

4 TTGGTTGGGCAACACA 19

THIS PAGE LEFT BLANK

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 6, 2006, 16:00:58 ; Search time 85 Seconds  
(without alignments)  
418.250 Million cell updates/sec

Title: US-10-671-074-172

Perfect score: 20

Sequence: 1 gcttgggtgggcaacacat 20

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2605320

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq:\*
- 7: /cgn2\_6/ptodata/1/ina/pp COMB.seq:\*
- 8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*
- 9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	20	100.0	254	US-09-794-928A-12
2	16	80.0	10603	US-09-949-016-16598
3	15	75.0	601	US-09-949-016-42568
4	15	75.0	1342	US-08-976-259-66
5	15	75.0	1342	US-09-956-004-66
6	15	75.0	1701	US-09-543-681A-3630
7	15	75.0	2181	US-09-328-352-3989
8	15	75.0	3593	US-09-644-934-10
9	15	75.0	5205	US-09-724-126A-18
10	15	75.0	6308	US-09-949-016-126A-1
11	15	75.0	222452	US-09-949-016-126A-1
12	15	75.0	1830121	US-09-557-884-1
13	15	75.0	1830121	US-09-643-990A-1
14	15	75.0	1830121	US-10-158-865-1
15	15	75.0	349	US-09-513-999C-26749
16	14	70.0	410	US-09-312-283C-77
17	14	70.0	440	US-09-188-930-77
18	14	70.0	601	US-09-949-016-27657
19	14	70.0	601	US-09-949-016-34062
20	14	70.0	601	US-09-949-016-158206
21	14	70.0	601	US-09-949-016-158206
22	14	70.0	601	US-09-949-016-158206
23	14	70.0	707	US-09-442-143A-7
24	14	70.0	707	US-09-902-563-7

25	14	70.0	1194	US-09-583-110-2269	Sequence 2269, Ap
26	14	70.0	1194	US-09-107-433-294	Sequence 294, Ap
c 27	14	70.0	1279	US-08-961-083-47	Sequence 47, Appl
c 28	14	70.0	1279	US-09-536-784-47	Sequence 47, Appl
c 29	14	70.0	1279	US-09-765-271-47	Sequence 47, Appl
c 30	14	70.0	1279	US-09-765-272A-47	Sequence 47, Appl
c 31	14	70.0	1284	US-09-196-857-1	Sequence 1, Appl
c 32	14	70.0	1284	US-09-583-110-1279	Sequence 1, Appl
c 33	14	70.0	1308	US-09-107-433-1653	Sequence 1279, Ap
c 34	14	70.0	1449	US-09-917-254-22	Sequence 1653, Ap
c 35	14	70.0	2184	US-09-949-016-4402	Sequence 22, Appl
c 36	14	70.0	2743	US-08-396-479B-3	Sequence 4402, Ap
c 37	14	70.0	2743	US-08-818-823-3	Sequence 3, Appl
c 38	14	70.0	2749	US-08-124-981A-1	Sequence 3, Appl
c 39	14	70.0	2751	US-09-037-190-45	Sequence 1, Appl
c 40	14	70.0	2751	US-09-037-192-45	Sequence 45, Appl
c 41	14	70.0	2751	US-09-037-192-45	Sequence 45, Appl
c 42	14	70.0	2751	US-09-037-143-45	Sequence 45, Appl
c 43	14	70.0	2751	US-09-049-691-45	Sequence 45, Appl
c 44	14	70.0	2751	US-08-260-174-45	Sequence 45, Appl
c 45	14	70.0	2751	US-09-338-128A-45	Sequence 45, Appl
c 46	14	70.0	2853	PCT-US94-07297-36	Sequence 45, Appl
c 47	14	70.0	2935	US-09-949-016-823	Sequence 36, Appl
c 48	14	70.0	3082	US-09-220-132-68	Sequence 823, Appl
c 49	14	70.0	4152	US-09-540-236-1660	Sequence 68, Appl
c 50	14	70.0	4464	US-09-949-016-5322	Sequence 1660, Ap
c 51	14	70.0	4539	US-09-704-725-1	Sequence 5322, Ap
c 52	14	70.0	5403	US-09-442-143A-3	Sequence 1, Appl
c 53	14	70.0	5403	US-09-902-563-3	Sequence 3, Appl
c 54	14	70.0	5502	US-09-232-346-45	Sequence 3, Appl
c 55	14	70.0	5502	US-10-040-430-45	Sequence 45, Appl
c 56	14	70.0	16995	US-08-961-527-82	Sequence 45, Appl
c 57	14	70.0	77036	US-09-949-016-16156	Sequence 82, Appl
c 58	14	70.0	101558	US-09-949-016-12243	Sequence 16156, A
c 59	14	70.0	136058	US-09-949-016-12243	Sequence 12243, A
c 60	14	70.0	136480	US-09-949-016-12565	Sequence 12565, A
c 61	14	70.0	167708	US-09-949-016-17064	Sequence 17064, A
c 62	14	70.0	191433	US-09-949-016-16423	Sequence 16423, A
c 63	14	70.0	278866	US-09-949-016-16144	Sequence 16144, A
c 64	14	70.0	278866	US-09-949-016-13922	Sequence 13922, A
c 65	14	70.0	278866	US-09-949-016-13924	Sequence 13923, A
c 66	14	70.0	278866	US-09-949-016-13925	Sequence 13924, A
c 67	14	70.0	278866	US-09-949-016-13925	Sequence 13925, A
c 68	14	70.0	278866	US-09-949-016-14699	Sequence 14699, A
c 69	14	70.0	278866	US-09-949-016-14700	Sequence 14700, A
c 70	14	70.0	278866	US-09-949-016-14701	Sequence 14701, A
c 71	14	70.0	278866	US-09-949-016-14702	Sequence 14702, A
c 72	14	70.0	278866	US-09-949-016-14703	Sequence 14703, A
c 73	13	65.0	24	US-09-009-443-5	Sequence 5, Appl
c 74	13	65.0	161	US-08-952-061-4	Sequence 4, Appl
c 75	13	65.0	216	US-09-084-303B-208	Sequence 208, Ap
c 76	13	65.0	216	US-09-621-976-14957	Sequence 14957, A
c 77	13	65.0	224	US-09-270-767-6633	Sequence 6633, Ap
c 78	13	65.0	224	US-09-270-767-21915	Sequence 21915, A
c 79	13	65.0	395	US-09-621-976-18231	Sequence 18231, A
c 80	13	65.0	400	US-08-956-171E-2971	Sequence 2971, Ap
c 81	13	65.0	400	US-08-781-986A-2971	Sequence 2971, Ap
c 82	13	65.0	408	US-08-953-171-4	Sequence 4, Appl
c 83	13	65.0	417	US-09-270-767-3143	Sequence 3143, Ap
c 84	13	65.0	417	US-09-270-767-18425	Sequence 18425, A
c 85	13	65.0	480	US-09-248-796A-4633	Sequence 4633, Ap
c 86	13	65.0	535	US-09-270-767-1931	Sequence 1931, Ap
c 87	13	65.0	535	US-09-270-767-17213	Sequence 17213, A
c 88	13	65.0	601	US-09-949-016-36146	Sequence 36146, A
c 89	13	65.0	601	US-09-949-016-36147	Sequence 36147, A
c 90	13	65.0	601	US-09-949-016-32706	Sequence 32706, A
c 91	13	65.0	601	US-09-949-016-71318	Sequence 71318, A
c 92	13	65.0	601	US-09-949-016-73923	Sequence 73923, A
c 93	13	65.0	601	US-09-949-016-74029	Sequence 74029, A
c 94	13	65.0	601	US-09-949-016-109498	Sequence 109498, A
c 95	13	65.0	601	US-09-949-016-134304	Sequence 134304, A
c 96	13	65.0	601	US-09-949-016-134305	Sequence 134305, A
c 97	13	65.0	601	US-09-949-016-139357	Sequence 139357, A

c 98	13	65.0	601	3	US-09-949-016-139358	Sequence 139358,	171	13	65.0	39154	3	US-09-949-016-12384	Sequence 12384, A
c 99	13	65.0	601	3	US-09-949-016-139359	Sequence 139359,	172	13	65.0	39154	3	US-09-949-016-12801	Sequence 12801, A
c 100	13	65.0	601	3	US-09-949-016-139360	Sequence 139360,	173	13	65.0	39443	3	US-09-949-016-14326	Sequence 14326, A
c 101	13	65.0	601	3	US-09-949-016-153559	Sequence 133559,	174	13	65.0	39443	3	US-09-949-016-14327	Sequence 14327, A
c 102	13	65.0	601	3	US-09-949-016-169479	Sequence 169479,	c 175	13	65.0	40123	3	US-09-949-002-646	Sequence 646, App
c 103	13	65.0	601	3	US-09-949-016-169480	Sequence 169480,	c 176	13	65.0	40124	3	US-09-949-002-775	Sequence 775, App
c 104	13	65.0	601	3	US-09-949-016-169481	Sequence 169481,	c 177	13	65.0	42348	3	US-09-949-016-17157	Sequence 17157, A
c 105	13	65.0	601	3	US-09-949-016-174750	Sequence 174750,	c 178	13	65.0	44971	3	US-09-949-016-17049	Sequence 17049, A
c 106	13	65.0	601	3	US-09-949-016-174751	Sequence 174751,	c 179	13	65.0	51508	3	US-09-949-016-16681	Sequence 16681, A
c 107	13	65.0	601	3	US-09-540-236-1853	Sequence 1853,	c 180	13	65.0	54420	3	US-09-949-016-15422	Sequence 15422, A
c 108	13	65.0	609	3	US-09-252-991A-8534	Sequence 8534, Ap	c 181	13	65.0	55841	3	US-09-949-016-16602	Sequence 16602, A
c 109	13	65.0	640	3	US-09-270-767-7378	Sequence 7378, Ap	c 182	13	65.0	56904	3	US-09-949-016-15501	Sequence 15501, A
c 110	13	65.0	640	3	US-09-270-767-7378	Sequence 7378, Ap	c 183	13	65.0	58879	3	US-09-949-016-16052	Sequence 16052, A
c 111	13	65.0	831	3	US-09-252-931A-8588	Sequence 8588, Ap	c 184	13	65.0	67581	3	US-09-949-016-14768	Sequence 14768, A
c 112	13	65.0	918	2	US-08-743-637B-162	Sequence 162, App	c 185	13	65.0	67643	3	US-09-949-016-14760	Sequence 14760, A
c 113	13	65.0	918	3	US-08-526-840B-162	Sequence 162, App	c 186	13	65.0	69813	3	US-09-949-016-12455	Sequence 12455, A
c 114	13	65.0	963	3	US-09-252-991A-7371	Sequence 7371, Ap	c 187	13	65.0	69813	3	US-09-949-016-13905	Sequence 13905, A
c 115	13	65.0	979	3	US-09-270-767-13232	Sequence 13232, A	c 188	13	65.0	69813	3	US-09-949-016-13906	Sequence 13906, A
c 116	13	65.0	1041	2	US-08-580-545B-5	Sequence 5, Appli	c 189	13	65.0	69833	3	US-09-949-016-12861	Sequence 12861, A
c 117	13	65.0	1041	3	US-08-262-653A-5	Sequence 5, Appli	c 190	13	65.0	80246	3	US-09-078-294-4	Sequence 4, Appli
c 118	13	65.0	1050	3	US-09-252-991A-7130	Sequence 7130, Ap	c 191	13	65.0	80269	3	US-09-949-016-15681	Sequence 15681, A
c 119	13	65.0	1131	3	US-09-540-236-260	Sequence 260, App	c 192	13	65.0	80595	3	US-09-078-294-3	Sequence 3, Appli
c 120	13	65.0	1161	3	US-08-956-171E-497	Sequence 497, App	c 193	13	65.0	95566	3	US-09-949-016-11877	Sequence 11877, A
c 121	13	65.0	1161	3	US-08-781-986A-497	Sequence 497, App	c 194	13	65.0	97423	3	US-09-949-016-12742	Sequence 12742, A
c 122	13	65.0	1181	3	US-09-303-518D-875	Sequence 875, App	c 195	13	65.0	97424	3	US-09-949-016-15576	Sequence 15576, A
c 123	13	65.0	1260	3	US-09-583-110-280	Sequence 280, App	c 196	13	65.0	99629	3	US-09-596-002-37	Sequence 37, Appl
c 124	13	65.0	1281	3	US-09-107-433-1849	Sequence 1849, Ap	c 197	13	65.0	100848	3	US-09-596-002-39	Sequence 39, Appl
c 125	13	65.0	1348	3	US-08-774-528-404	Sequence 404, App	c 198	13	65.0	108440	3	US-09-949-016-12065	Sequence 12065, A
c 126	13	65.0	1348	3	US-10-120-988-404	Sequence 404, App	c 199	13	65.0	108441	3	US-09-949-016-14090	Sequence 14090, A
c 127	13	65.0	1407	3	US-09-540-236-740	Sequence 740, App	c 200	13	65.0	109250	3	US-09-949-016-12530	Sequence 12530, A
c 128	13	65.0	1475	3	US-08-889-841B-30	Sequence 30, Appl	c 201	13	65.0	109251	3	US-09-949-016-17321	Sequence 17321, A
c 129	13	65.0	1475	3	US-08-889-841B-32	Sequence 32, Appl	c 202	13	65.0	112465	3	US-09-949-016-16424	Sequence 16424, A
c 130	13	65.0	1475	3	US-09-419-362-30	Sequence 30, Appl	c 203	13	65.0	1252132	3	US-09-949-016-13845	Sequence 13845, A
c 131	13	65.0	1475	3	US-09-419-362-32	Sequence 32, Appl	c 204	13	65.0	1252145	3	US-09-949-016-12371	Sequence 12371, A
c 132	13	65.0	1656	3	US-09-248-796A-3263	Sequence 3263, Ap	c 205	13	65.0	155617	3	US-09-949-016-16191	Sequence 16191, A
c 133	13	65.0	1659	3	US-09-107-532A-1211	Sequence 1211, Ap	c 206	13	65.0	160759	3	US-09-949-016-16514	Sequence 16514, A
c 134	13	65.0	1842	3	US-09-328-352-940	Sequence 940, App	c 207	13	65.0	177797	3	US-09-949-016-14125	Sequence 14125, A
c 135	13	65.0	1860	3	US-09-248-796A-1590	Sequence 1590, Ap	c 208	13	65.0	187580	3	US-09-949-016-13266	Sequence 13266, A
c 136	13	65.0	1995	3	US-09-771-161A-31	Sequence 31, Appl	c 209	13	65.0	221545	3	US-09-949-016-13875	Sequence 13875, A
c 137	13	65.0	2311	3	US-08-800-729-66	Sequence 66, Appl	c 210	13	65.0	227979	3	US-09-949-016-11842	Sequence 11842, A
c 138	13	65.0	2388	3	US-09-205-258-154	Sequence 154, App	c 211	13	65.0	229354	3	US-09-705-400-64	Sequence 64, Appl
c 139	13	65.0	2388	3	US-10-004-860-154	Sequence 154, App	c 212	13	65.0	237241	3	US-09-949-016-16101	Sequence 16101, A
c 140	13	65.0	2394	3	US-09-800-729-33	Sequence 33, Appl	c 213	13	65.0	247299	3	US-09-949-016-17530	Sequence 17590, A
c 141	13	65.0	2502	3	US-09-252-991A-8656	Sequence 8656, Ap	c 214	13	65.0	340380	3	US-09-949-016-14179	Sequence 14179, A
c 142	13	65.0	2631	3	US-09-487-558B-43	Sequence 43, Appl	c 215	13	65.0	818128	3	US-09-949-016-14546	Sequence 14546, A
c 143	13	65.0	2667	3	US-08-303-518D-877	Sequence 877, App	c 216	13	65.0	818128	3	US-09-949-016-14547	Sequence 14547, A
c 144	13	65.0	2769	3	US-09-303-518D-883	Sequence 883, App	c 217	13	65.0	818128	3	US-09-949-016-14548	Sequence 14548, A
c 145	13	65.0	2800	3	US-08-714-918-35	Sequence 35, Appl	c 218	13	65.0	818128	3	US-09-949-016-14549	Sequence 14549, A
c 146	13	65.0	2800	3	US-09-265-315-35	Sequence 35, Appl	c 219	13	65.0	818128	3	US-09-949-016-14550	Sequence 14550, A
c 147	13	65.0	2800	3	US-09-266-417-35	Sequence 35, Appl	c 220	13	65.0	818128	3	US-09-949-016-14551	Sequence 14551, A
c 148	13	65.0	2800	3	US-09-266-417-35	Sequence 35, Appl	c 221	13	65.0	818128	3	US-09-949-016-14552	Sequence 14552, A
c 149	13	65.0	2800	3	US-09-528-709-35	Sequence 35, Appl	c 222	13	65.0	818128	3	US-09-949-016-14553	Sequence 14553, A
c 150	13	65.0	2800	3	US-09-527-745-35	Sequence 35, Appl	c 223	13	65.0	818128	3	US-09-949-016-14554	Sequence 14554, A
c 151	13	65.0	4234	3	US-09-976-594-513	Sequence 513, App	c 224	13	65.0	818128	3	US-09-949-016-14555	Sequence 14555, A
c 152	13	65.0	4826	3	US-08-956-171E-275	Sequence 275, App	c 225	13	65.0	818128	3	US-09-949-016-14556	Sequence 14556, A
c 153	13	65.0	4826	3	US-08-781-986A-275	Sequence 275, App	c 226	13	65.0	818128	3	US-09-949-016-14557	Sequence 14557, A
c 154	13	65.0	5162	3	US-09-298-367B-1	Sequence 1, Appli	c 227	13	65.0	818128	3	US-09-949-016-14558	Sequence 14558, A
c 155	13	65.0	5162	3	US-09-298-367B-4	Sequence 4, Appli	c 228	13	65.0	818128	3	US-09-949-016-14559	Sequence 14559, A
c 156	13	65.0	5262	3	US-09-298-367B-5	Sequence 5, Appli	c 229	13	65.0	818128	3	US-09-949-016-14560	Sequence 14560, A
c 157	13	65.0	6360	3	US-09-171-699-9	Sequence 9, Appli	c 230	13	65.0	818128	3	US-09-949-016-14561	Sequence 14561, A
c 158	13	65.0	6850	3	US-09-298-367B-2	Sequence 2, Appli	c 231	13	65.0	818128	3	US-09-949-016-14562	Sequence 14562, A
c 159	13	65.0	7159	3	US-08-956-171E-302	Sequence 302, App	c 232	13	65.0	818128	3	US-09-949-016-14564	Sequence 14564, A
c 160	13	65.0	8035	3	US-08-781-986A-302	Sequence 302, App	c 233	13	65.0	818128	3	US-09-949-016-14565	Sequence 14565, A
c 161	13	65.0	8617	3	US-09-949-016-14727	Sequence 14727, Ap	c 234	13	65.0	818128	3	US-09-949-016-14566	Sequence 14566, A
c 162	13	65.0	8617	3	US-09-949-016-14663	Sequence 14663, A	c 235	13	65.0	818128	3	US-09-949-016-14567	Sequence 14567, A
c 163	13	65.0	8618	3	US-09-949-016-14286	Sequence 14286, A	c 236	12	60.0	24	3	US-09-482-180A-9	Sequence 9, Appli
c 164	13	65.0	14012	3	US-09-949-016-14625	Sequence 14625, A	c 237	12	60.0	36	3	US-09-569-601A-2	Sequence 2, Appli
c 165	13	65.0	15695	3	US-09-949-016-15644	Sequence 15644, A	c 238	12	60.0	36	3	US-09-569-601A-18	Sequence 18, Appli
c 166	13	65.0	28321	3	US-09-949-016-11936	Sequence 11936, A	c 239	12	60.0	133	3	US-09-513-999C-9885	Sequence 9885, Ap
c 167	13	65.0	28325	3	US-09-949-016-16622	Sequence 16622, A	c 240	12	60.0	136	3	US-09-471-276-704	Sequence 704, App
c 168	13	65.0	32099	3	US-09-949-016-16562	Sequence 16562, A	c 241	12	60.0	202	3	US-09-513-999C-23748	Sequence 23748, A
c 169	13	65.0	36855	3	US-09-949-016-17095	Sequence 17095, A	c 242	12	60.0	212	3	US-09-513-999C-18559	Sequence 18559, A
c 170	13	65.0	37613	3	US-09-949-016-12669	Sequence 12669, A	c 243	12	60.0	227	3	US-09-754-066-19	Sequence 19, Appli



244	12	60.0	245	3	US-09-602-787A-649	Sequence 649, App	317	12	60.0	601	3	US-09-949-016-56995	Sequence 56995, A
245	12	60.0	245	3	US-09-513-999C-10946	Sequence 10946, A	318	12	60.0	601	3	US-09-949-016-57074	Sequence 57074, A
246	12	60.0	247	3	US-09-513-999C-24147	Sequence 24147, A	319	12	60.0	601	3	US-09-949-016-57075	Sequence 57075, A
247	12	60.0	253	3	US-09-513-999C-25936	Sequence 25936, A	320	12	60.0	601	3	US-09-949-016-59627	Sequence 59627, A
248	12	60.0	266	3	US-09-513-294A-3633	Sequence 3633, Ap	c 321	12	60.0	601	3	US-09-949-016-60171	Sequence 60171, A
249	12	60.0	291	3	US-08-778-717-1	Sequence 1, Appli	c 322	12	60.0	601	3	US-09-949-016-60328	Sequence 60328, A
250	12	60.0	303	3	US-09-540-236-825	Sequence 825, App	323	12	60.0	601	3	US-09-949-016-62539	Sequence 62539, A
251	12	60.0	336	3	US-09-513-999C-9569	Sequence 9569, Ap	324	12	60.0	601	3	US-09-949-016-62540	Sequence 62540, A
252	12	60.0	343	3	US-09-513-559-3747	Sequence 3747, Ap	325	12	60.0	601	3	US-09-949-016-62541	Sequence 62541, A
253	12	60.0	359	3	US-09-270-767-2723	Sequence 2723, Ap	326	12	60.0	601	3	US-09-949-016-62542	Sequence 62542, A
254	12	60.0	359	3	US-09-270-767-18005	Sequence 18005, A	327	12	60.0	601	3	US-09-949-016-62543	Sequence 62543, A
255	12	60.0	361	3	US-09-513-999C-32924	Sequence 32924, A	328	12	60.0	601	3	US-09-949-016-62867	Sequence 62867, A
256	12	60.0	365	3	US-09-513-999C-33710	Sequence 33710, A	329	12	60.0	601	3	US-09-949-016-63236	Sequence 63236, A
257	12	60.0	390	3	US-09-711-164-216	Sequence 216, App	c 330	12	60.0	601	3	US-09-949-016-67772	Sequence 67772, A
258	12	60.0	390	3	US-09-711-164-248	Sequence 248, App	331	12	60.0	601	3	US-09-949-016-68607	Sequence 68607, A
259	12	60.0	393	3	US-09-489-038A-3764	Sequence 3764, Ap	332	12	60.0	601	3	US-09-949-016-69640	Sequence 69640, A
260	12	60.0	399	3	US-09-270-767-1659	Sequence 1659, Ap	c 333	12	60.0	601	3	US-09-949-016-69812	Sequence 69812, A
261	12	60.0	399	3	US-09-270-767-16941	Sequence 16941, A	c 334	12	60.0	601	3	US-09-949-016-76583	Sequence 76583, A
262	12	60.0	423	3	US-09-328-352-3580	Sequence 3580, Ap	c 335	12	60.0	601	3	US-09-949-016-76583	Sequence 76583, A
263	12	60.0	428	3	US-09-621-976-17842	Sequence 17842, A	c 336	12	60.0	601	3	US-09-949-016-78030	Sequence 78030, A
264	12	60.0	462	3	US-09-621-976-3261	Sequence 3261, Ap	c 337	12	60.0	601	3	US-09-949-016-80199	Sequence 80199, A
265	12	60.0	467	3	US-09-270-767-2340	Sequence 2340, Ap	338	12	60.0	601	3	US-09-949-016-80726	Sequence 80726, A
266	12	60.0	467	3	US-09-270-767-17822	Sequence 17822, A	339	12	60.0	601	3	US-09-949-016-80728	Sequence 80728, A
267	12	60.0	471	2	US-08-599-602-1	Sequence 1, Appli	340	12	60.0	601	3	US-09-949-016-83757	Sequence 83757, A
268	12	60.0	471	2	US-09-197-816-1	Sequence 1, Appli	c 341	12	60.0	601	3	US-09-949-016-83757	Sequence 83757, A
269	12	60.0	486	3	US-09-252-991A-11364	Sequence 11364, A	c 342	12	60.0	601	3	US-09-949-016-87362	Sequence 87362, A
270	12	60.0	487	3	US-09-621-976-10207	Sequence 10207, A	c 343	12	60.0	601	3	US-09-949-016-91374	Sequence 91374, A
271	12	60.0	491	3	US-09-621-976-17740	Sequence 17740, A	c 344	12	60.0	601	3	US-09-949-016-92943	Sequence 92943, A
272	12	60.0	500	3	US-09-270-767-5759	Sequence 5759, Ap	c 345	12	60.0	601	3	US-09-949-016-107318	Sequence 107318, A
273	12	60.0	500	3	US-09-270-767-21041	Sequence 21041, A	c 346	12	60.0	601	3	US-09-949-016-107422	Sequence 107422, A
274	12	60.0	505	3	US-09-270-767-6920	Sequence 6920, Ap	c 347	12	60.0	601	3	US-09-949-016-107423	Sequence 107423, A
275	12	60.0	505	3	US-09-270-767-22202	Sequence 22202, A	c 348	12	60.0	601	3	US-09-949-016-107491	Sequence 107491, A
276	12	60.0	508	3	US-09-058-389A-21	Sequence 21, Appli	349	12	60.0	601	3	US-09-949-016-107492	Sequence 107492, A
277	12	60.0	508	3	US-09-611-781-21	Sequence 21, Appli	c 350	12	60.0	601	3	US-09-949-016-108568	Sequence 108568, A
278	12	60.0	509	3	US-09-270-767-8814	Sequence 8814, Ap	c 351	12	60.0	601	3	US-09-949-016-108569	Sequence 108569, A
279	12	60.0	509	3	US-09-270-767-24096	Sequence 24096, A	c 352	12	60.0	601	3	US-09-949-016-117608	Sequence 117608, A
280	12	60.0	526	3	US-09-602-777A-121	Sequence 121, App	c 353	12	60.0	601	3	US-09-949-016-119478	Sequence 119478, A
281	12	60.0	534	3	US-09-621-976-17102	Sequence 17102, A	c 354	12	60.0	601	3	US-09-949-016-127612	Sequence 127612, A
282	12	60.0	538	3	US-09-621-976-584	Sequence 584, App	c 355	12	60.0	601	3	US-09-949-016-127615	Sequence 127615, A
283	12	60.0	549	3	US-09-533-559-1306	Sequence 1306, Ap	c 356	12	60.0	601	3	US-09-949-016-127615	Sequence 127615, A
284	12	60.0	577	3	US-09-270-767-7037	Sequence 7037, Ap	c 357	12	60.0	601	3	US-09-949-016-127949	Sequence 127949, A
285	12	60.0	579	3	US-09-270-767-22319	Sequence 22319, A	c 358	12	60.0	601	3	US-09-949-016-127951	Sequence 127951, A
286	12	60.0	579	3	US-08-953-171-3	Sequence 3, Appli	c 359	12	60.0	601	3	US-09-949-016-127952	Sequence 127952, A
287	12	60.0	579	3	US-09-540-236-1408	Sequence 1408, Ap	c 360	12	60.0	601	3	US-09-949-016-133403	Sequence 133403, A
288	12	60.0	595	2	US-08-076-091C-5	Sequence 5, Appli	c 361	12	60.0	601	3	US-09-949-016-137820	Sequence 137820, A
289	12	60.0	595	2	US-08-285-641-5	Sequence 5, Appli	c 362	12	60.0	601	3	US-09-949-016-141122	Sequence 141122, A
290	12	60.0	601	3	US-09-949-016-19927	Sequence 19927, A	c 363	12	60.0	601	3	US-09-949-016-143169	Sequence 143169, A
291	12	60.0	601	3	US-09-949-016-19929	Sequence 19929, A	c 364	12	60.0	601	3	US-09-949-016-148693	Sequence 148693, A
292	12	60.0	601	3	US-09-949-016-19930	Sequence 19930, A	c 365	12	60.0	601	3	US-09-949-016-150176	Sequence 150176, A
293	12	60.0	601	3	US-09-949-016-19931	Sequence 19931, A	c 366	12	60.0	601	3	US-09-949-016-151566	Sequence 151566, A
294	12	60.0	601	3	US-09-949-016-21468	Sequence 21468, A	c 367	12	60.0	601	3	US-09-949-016-152454	Sequence 152454, A
295	12	60.0	601	3	US-09-949-016-21839	Sequence 21839, A	c 368	12	60.0	601	3	US-09-949-016-156522	Sequence 156522, A
296	12	60.0	601	3	US-09-949-016-22278	Sequence 22278, A	c 369	12	60.0	601	3	US-09-949-016-156985	Sequence 156985, A
297	12	60.0	601	3	US-09-949-016-25089	Sequence 25089, A	c 370	12	60.0	601	3	US-09-949-016-157031	Sequence 157031, A
298	12	60.0	601	3	US-09-949-016-25090	Sequence 25090, A	c 371	12	60.0	601	3	US-09-949-016-158153	Sequence 158153, A
299	12	60.0	601	3	US-09-949-016-25930	Sequence 25930, A	c 372	12	60.0	601	3	US-09-949-016-160326	Sequence 160326, A
300	12	60.0	601	3	US-09-949-016-30206	Sequence 30206, A	c 373	12	60.0	601	3	US-09-949-016-160327	Sequence 160327, A
301	12	60.0	601	3	US-09-949-016-35483	Sequence 35483, A	c 374	12	60.0	601	3	US-09-949-016-160328	Sequence 160328, A
302	12	60.0	601	3	US-09-949-016-35812	Sequence 35812, A	c 375	12	60.0	601	3	US-09-949-016-160329	Sequence 160329, A
303	12	60.0	601	3	US-09-949-016-35890	Sequence 35890, A	c 376	12	60.0	601	3	US-09-949-016-160330	Sequence 160330, A
304	12	60.0	601	3	US-09-949-016-35967	Sequence 35967, A	c 377	12	60.0	601	3	US-09-949-016-160702	Sequence 160702, A
305	12	60.0	601	3	US-09-949-016-36054	Sequence 36054, A	c 378	12	60.0	601	3	US-09-949-016-160703	Sequence 160703, A
306	12	60.0	601	3	US-09-949-016-46189	Sequence 46189, A	c 379	12	60.0	601	3	US-09-949-016-164880	Sequence 164880, A
307	12	60.0	601	3	US-09-949-016-46191	Sequence 46191, A	c 380	12	60.0	601	3	US-09-949-016-166838	Sequence 166838, A
308	12	60.0	601	3	US-09-949-016-46192	Sequence 46192, A	c 381	12	60.0	601	3	US-09-949-016-166839	Sequence 166839, A
309	12	60.0	601	3	US-09-949-016-46193	Sequence 46193, A	c 382	12	60.0	601	3	US-09-949-016-175881	Sequence 175881, A
310	12	60.0	601	3	US-09-949-016-52302	Sequence 52302, A	c 383	12	60.0	601	3	US-09-949-016-178108	Sequence 178108, A
311	12	60.0	601	3	US-09-949-016-52562	Sequence 52562, A	c 384	12	60.0	601	3	US-09-949-016-178109	Sequence 178109, A
312	12	60.0	601	3	US-09-949-016-52563	Sequence 52563, A	c 385	12	60.0	601	3	US-09-949-016-180528	Sequence 180528, A
313	12	60.0	601	3	US-09-949-016-55967	Sequence 55967, A	c 386	12	60.0	601	3	US-09-949-016-180731	Sequence 180731, A
314	12	60.0	601	3	US-09-949-016-56550	Sequence 56550, A	c 387	12	60.0	601	3	US-09-949-016-180732	Sequence 180732, A
315	12	60.0	601	3	US-09-949-016-56551	Sequence 56551, A	c 388	12	60.0	601	3	US-09-949-016-194980	Sequence 194980, A
316	12	60.0	601	3	US-09-949-016-56552	Sequence 56552, A	c 389	12	60.0	601	3	US-09-949-016-194981	Sequence 194981, A
												US-09-949-016-197130	Sequence 197130, A

C 390	12	60.0	601	3	US-09-949-016-197248	Sequence 197248,	463	12	60.0	1351	3	US-08-796-792-3	Sequence 3, Appli
C 391	12	60.0	601	3	US-09-949-016-197366	Sequence 197366,	464	12	60.0	1351	3	US-09-491-795-3	Sequence 3, Appli
C 392	12	60.0	601	3	US-09-949-016-197484	Sequence 197484,	465	12	60.0	1398	3	US-09-328-352-3200	Sequence 3200, Ap
C 393	12	60.0	601	3	US-09-949-016-199185	Sequence 199185,	466	12	60.0	1413	3	US-09-543-681A-3221	Sequence 3221, Ap
C 394	12	60.0	601	3	US-09-949-016-199906	Sequence 199906,	C 467	12	60.0	1419	3	US-09-540-236-1840	Sequence 1840, Ap
C 395	12	60.0	601	3	US-09-949-016-200360	Sequence 200360,	468	12	60.0	1420	3	US-09-482-180A-1	Sequence 1, Appli
C 396	12	60.0	601	3	US-09-949-016-200361	Sequence 200361,	C 469	12	60.0	1460	3	US-09-573-080A-364	Sequence 364, App
C 397	12	60.0	601	3	US-09-949-016-202637	Sequence 202637,	C 470	12	60.0	1605	3	US-09-252-991A-15515	Sequence 15515, A
C 398	12	60.0	601	3	US-09-949-002-2264	Sequence 2264, Ap	471	12	60.0	1701	2	US-08-484-993B-40	Sequence 40, Appl
C 399	12	60.0	601	3	US-09-949-002-2265	Sequence 2265, Ap	472	12	60.0	1701	2	US-08-484-158B-40	Sequence 40, Appl
C 400	12	60.0	601	3	US-09-949-002-2266	Sequence 2266, Ap	473	12	60.0	1701	2	US-08-484-596A-40	Sequence 40, Appl
C 401	12	60.0	601	3	US-09-949-002-8414	Sequence 8414, Ap	474	12	60.0	1701	2	US-08-480-150A-40	Sequence 40, Appl
C 402	12	60.0	601	3	US-09-949-002-8415	Sequence 8415, Ap	475	12	60.0	1701	2	US-08-458-731-40	Sequence 40, Appl
C 403	12	60.0	601	3	US-09-949-002-8416	Sequence 8416, Ap	476	12	60.0	1701	3	US-08-149-223A-40	Sequence 40, Appl
C 404	12	60.0	601	3	US-09-949-002-9535	Sequence 9535, Ap	C 477	12	60.0	1704	3	US-09-711-164-203	Sequence 203, App
C 405	12	60.0	601	3	US-09-949-002-9536	Sequence 9536, Ap	C 478	12	60.0	1719	3	US-09-248-796A-4662	Sequence 4662, Ap
C 406	12	60.0	601	3	US-09-949-002-9537	Sequence 9537, Ap	C 479	12	60.0	1759	3	US-10-104-047-1283	Sequence 1283, Ap
C 407	12	60.0	603	2	US-08-799-464A-21	Sequence 21, Appl	C 480	12	60.0	1788	3	US-09-404-296B-7	Sequence 7, Appli
C 408	12	60.0	603	6	PCT-US95-09927-21	Sequence 12, Appl	C 481	12	60.0	1801	2	US-09-130-114-3	Sequence 3, Appli
C 409	12	60.0	606	2	US-08-131-625B-12	Sequence 12, Appl	C 482	12	60.0	1845	3	US-08-476-102A-7	Sequence 7, Appli
C 410	12	60.0	606	3	US-08-855-531D-11	Sequence 11, Appl	483	12	60.0	1863	3	US-09-902-540-4947	Sequence 4947, Ap
C 411	12	60.0	606	3	US-08-855-526B-11	Sequence 11, Appl	484	12	60.0	1878	3	US-09-489-039A-2608	Sequence 2608, Ap
C 412	12	60.0	606	3	US-08-301-435-20	Sequence 20, Appl	C 485	12	60.0	1913	3	US-09-533-559-757	Sequence 757, App
C 413	12	60.0	606	6	PCT-US95-10904-20	Sequence 20, Appl	C 486	12	60.0	1936	3	US-09-249-585A-1	Sequence 1, Appli
C 414	12	60.0	612	3	US-09-280-116-218	Sequence 218, App	C 487	12	60.0	1937	3	US-09-050-863-1	Sequence 1, Appli
C 415	12	60.0	612	3	US-09-385-982-215	Sequence 215, App	C 488	12	60.0	1937	3	US-09-359-081-1	Sequence 1, Appli
C 416	12	60.0	618	3	US-09-270-767-8516	Sequence 8516, Ap	489	12	60.0	1944	2	US-08-392-806A-3	Sequence 3, Appli
C 417	12	60.0	618	3	US-09-270-767-23798	Sequence 23798, A	490	12	60.0	1944	3	US-09-257-490-3	Sequence 3, Appli
C 418	12	60.0	636	3	US-09-710-279-743	Sequence 743, App	491	12	60.0	2086	3	US-09-540-236-1823	Sequence 1823, Ap
C 419	12	60.0	636	3	US-09-533-559-6334	Sequence 6334, Ap	C 492	12	60.0	2036	3	US-10-104-047-1669	Sequence 1669, Ap
C 420	12	60.0	645	3	US-09-134-001C-538	Sequence 538, App	493	12	60.0	2145	3	US-09-712-363-3	Sequence 3, Appli
C 421	12	60.0	656	3	US-09-533-559-1100	Sequence 1100, Ap	C 494	12	60.0	2151	3	US-09-270-767-11481	Sequence 11481, A
C 422	12	60.0	662	3	US-09-669-751-49	Sequence 49, Appl	C 495	12	60.0	2156	2	US-08-321-356-1	Sequence 1, Appli
C 423	12	60.0	668	3	US-09-270-767-8513	Sequence 8513, Ap	C 496	12	60.0	2156	2	US-08-321-356-3	Sequence 3, Appli
C 424	12	60.0	668	3	US-09-270-767-23795	Sequence 23795, A	C 497	12	60.0	2172	3	US-09-543-681A-1285	Sequence 1285, Ap
C 425	12	60.0	672	3	US-09-280-116-217	Sequence 217, App	C 498	12	60.0	2175	3	US-10-101-464A-838	Sequence 838, App
C 426	12	60.0	672	3	US-08-778-717-6	Sequence 6, Appli	C 499	12	60.0	2241	3	US-09-710-279-651	Sequence 651, App
C 427	12	60.0	672	3	US-08-778-717-8	Sequence 8, Appli	500	12	60.0	2250	3	US-09-540-236-130	Sequence 130, App
C 428	12	60.0	690	3	US-08-778-717-18	Sequence 18, Appli	501	12	60.0	2256	3	US-08-969-415-1	Sequence 1, Appli
C 429	12	60.0	699	3	US-09-540-236-141	Sequence 141, App	502	12	60.0	2265	3	US-09-270-767-14284	Sequence 14284, A
C 430	12	60.0	702	3	US-09-248-796A-3522	Sequence 3522, Ap	C 503	12	60.0	2297	3	US-09-471-016-12	Sequence 12, Appl
C 431	12	60.0	726	3	US-09-540-236-432	Sequence 432, App	C 504	12	60.0	2304	3	US-10-104-047-294	Sequence 294, App
C 432	12	60.0	795	3	US-09-248-796A-1124	Sequence 1124, Ap	C 505	12	60.0	2312	3	US-10-104-047-488	Sequence 488, App
C 433	12	60.0	801	3	US-09-269-446D-38	Sequence 38, Appl	506	12	60.0	2322	3	US-09-328-352-4077	Sequence 4077, Ap
C 434	12	60.0	810	3	US-08-778-717-10	Sequence 10, Appl	507	12	60.0	2337	3	US-09-134-001C-1031	Sequence 1031, Ap
C 435	12	60.0	815	3	US-08-778-717-12	Sequence 12, Appl	508	12	60.0	2352	3	US-09-328-352-3064	Sequence 3064, Ap
C 436	12	60.0	819	3	US-09-328-352-3246	Sequence 3246, Ap	509	12	60.0	2381	3	US-10-104-047-1619	Sequence 1619, Ap
C 437	12	60.0	837	3	US-09-134-000C-201	Sequence 201, App	C 510	12	60.0	2396	3	US-09-058-389A-10	Sequence 10, Appl
C 438	12	60.0	866	2	US-08-484-993B-46	Sequence 46, Appl	C 511	12	60.0	2396	3	US-09-611-781-10	Sequence 10, Appl
C 439	12	60.0	866	2	US-08-484-158B-46	Sequence 46, Appl	C 512	12	60.0	2435	3	US-09-306-593-1	Sequence 1, Appli
C 440	12	60.0	866	2	US-08-484-596A-46	Sequence 46, Appl	513	12	60.0	2466	3	US-09-363-243-1	Sequence 1, Appli
C 441	12	60.0	866	2	US-08-480-150A-46	Sequence 46, Appl	514	12	60.0	2621	3	US-08-817-441-63	Sequence 63, Appl
C 442	12	60.0	866	3	US-08-458-731-46	Sequence 46, Appl	515	12	60.0	2649	3	US-10-104-047-847	Sequence 847, App
C 443	12	60.0	866	3	US-08-149-223A-46	Sequence 46, Appl	516	12	60.0	2656	3	US-09-949-016-4867	Sequence 4867, Ap
C 444	12	60.0	870	3	US-09-710-279-63	Sequence 63, Appl	517	12	60.0	2682	3	US-09-614-221A-222	Sequence 222, App
C 445	12	60.0	989	3	US-09-671-317-283	Sequence 283, App	C 518	12	60.0	2742	3	US-09-328-352-2898	Sequence 2898, Ap
C 446	12	60.0	1020	3	US-09-477-135A-120	Sequence 120, App	C 519	12	60.0	2742	3	US-09-949-016-4295	Sequence 4295, Ap
C 447	12	60.0	1074	3	US-09-807-757C-5	Sequence 5, Appli	C 520	12	60.0	2743	2	US-08-317-707-1	Sequence 1, Appli
C 448	12	60.0	1083	3	US-09-602-777A-259	Sequence 259, App	521	12	60.0	2754	3	US-10-104-047-954	Sequence 954, App
C 449	12	60.0	1101	3	US-09-328-352-910	Sequence 910, App	C 522	12	60.0	2797	3	US-09-949-016-83	Sequence 83, Appl
C 450	12	60.0	1179	3	US-09-710-279-3305	Sequence 3305, Ap	C 523	12	60.0	2821	3	US-10-104-047-565	Sequence 565, App
C 451	12	60.0	1194	3	US-09-543-681A-1349	Sequence 1349, Ap	C 524	12	60.0	2887	3	US-09-949-016-3493	Sequence 3493, Ap
C 452	12	60.0	1197	3	US-09-134-001C-1516	Sequence 1516, Ap	C 525	12	60.0	2948	3	US-09-949-016-1106	Sequence 1106, Ap
C 453	12	60.0	1211	3	US-09-270-767-30424	Sequence 30424, A	C 526	12	60.0	2949	3	US-09-949-016-4761	Sequence 4761, Ap
C 454	12	60.0	1218	3	US-10-012-231A-5	Sequence 5, Appli	C 527	12	60.0	3032	3	US-09-710-279-3924	Sequence 3924, Ap
C 455	12	60.0	1218	3	US-10-015-389A-5	Sequence 5, Appli	528	12	60.0	3137	3	US-09-710-279-3842	Sequence 3842, Ap
C 456	12	60.0	1218	3	US-10-006-768A-5	Sequence 5, Appli	529	12	60.0	3141	2	US-08-658-665-66	Sequence 66, Appl
C 457	12	60.0	1218	3	US-10-015-671A-5	Sequence 5, Appli	530	12	60.0	3141	3	US-08-796-101-30	Sequence 30, Appl
C 458	12	60.0	1218	3	US-10-015-393A-5	Sequence 5, Appli	531	12	60.0	3141	3	US-09-085-273-66	Sequence 66, Appl
C 459	12	60.0	1218	3	US-10-011-833A-5	Sequence 5, Appli	532	12	60.0	3141	3	US-09-916-963-66	Sequence 66, Appl
C 460	12	60.0	1218	3	US-10-006-041A-5	Sequence 5, Appli	533	12	60.0	3370	3	US-09-710-279-4190	Sequence 4190, Ap
C 461	12	60.0	1218	3	US-10-012-064A-5	Sequence 5, Appli	534	12	60.0	3517	3	US-09-221-017B-281	Sequence 281, App
C 462	12	60.0	1227	3	US-09-252-991A-15322	Sequence 15322, A	C 535	12	60.0	3531	3	US-09-949-016-2301	Sequence 2301, Ap

C 536	12	60.0	3567	2	US-08-658-665-69	Sequence 69, Appl	C 609	12	60.0	9737	3	US-09-481-355-22	Sequence 22, Appl
C 537	12	60.0	3567	3	US-08-796-101-33	Sequence 33, Appl	C 610	12	60.0	9737	3	US-09-481-355-23	Sequence 23, Appl
C 538	12	60.0	3567	3	US-09-085-273-69	Sequence 69, Appl	C 611	12	60.0	9737	3	US-09-481-355-28	Sequence 28, Appl
C 539	12	60.0	3567	3	US-09-916-963-69	Sequence 69, Appl	C 612	12	60.0	9737	3	US-09-481-282-22	Sequence 22, Appl
C 540	12	60.0	3569	3	US-09-710-279-4443	Sequence 4443, Ap	C 613	12	60.0	9737	3	US-09-481-282-23	Sequence 23, Appl
C 541	12	60.0	3591	3	US-09-614-221A-336	Sequence 336, Appl	C 614	12	60.0	9737	3	US-09-481-282-28	Sequence 28, Appl
C 542	12	60.0	3592	3	US-09-775-398-84	Sequence 84, Appl	C 615	12	60.0	9737	3	US-09-455-659A-22	Sequence 22, Appl
C 543	12	60.0	3596	3	US-09-949-016-2135	Sequence 2135, Appl	C 616	12	60.0	9737	3	US-09-455-659A-23	Sequence 23, Appl
C 544	12	60.0	3716	3	US-09-999-843A-210	Sequence 210, Appl	C 617	12	60.0	9737	3	US-09-455-659A-28	Sequence 28, Appl
C 545	12	60.0	3716	3	US-10-020-445A-210	Sequence 210, Appl	C 618	12	60.0	9737	3	US-09-484-996-22	Sequence 22, Appl
C 546	12	60.0	3764	3	US-09-949-016-4511	Sequence 4511, Ap	C 619	12	60.0	9737	3	US-09-484-996-23	Sequence 23, Appl
C 547	12	60.0	3816	3	US-09-540-236-1820	Sequence 1820, Ap	C 620	12	60.0	9737	3	US-09-484-996-28	Sequence 28, Appl
C 548	12	60.0	3845	3	US-09-620-312D-554	Sequence 554, Appl	C 621	12	60.0	9737	3	US-09-479-123-22	Sequence 22, Appl
C 549	12	60.0	3871	3	US-09-949-016-631	Sequence 631, Appl	C 622	12	60.0	9737	3	US-09-479-123-23	Sequence 23, Appl
C 550	12	60.0	3941	3	US-09-408-865-2	Sequence 2, Appl1	C 623	12	60.0	9737	3	US-09-479-123-28	Sequence 28, Appl
C 551	12	60.0	4010	3	US-09-710-279-3540	Sequence 3540, Ap	C 624	12	60.0	9737	3	US-09-484-317A-22	Sequence 22, Appl
C 552	12	60.0	4019	3	US-09-710-279-4210	Sequence 4210, Ap	C 625	12	60.0	9737	3	US-09-484-317A-23	Sequence 23, Appl
C 553	12	60.0	4032	2	US-08-107-748-3	Sequence 3, Appl1	C 626	12	60.0	9737	3	US-09-484-317A-28	Sequence 28, Appl
C 554	12	60.0	4032	2	US-08-245-809-4	Sequence 4, Appl1	C 627	12	60.0	9737	3	US-09-484-317A-28	Sequence 28, Appl
C 555	12	60.0	4032	6	PCT-US92-01385-3	Sequence 3, Appl1	C 628	12	60.0	9737	3	US-09-276-820A-23	Sequence 23, Appl
C 556	12	60.0	4036	3	US-09-710-279-4287	Sequence 4287, Ap	C 629	12	60.0	9737	3	US-09-276-820A-23	Sequence 23, Appl
C 557	12	60.0	4075	2	US-08-658-665-67	Sequence 67, Appl	C 630	12	60.0	9837	3	US-09-221-017B-636	Sequence 636, App
C 558	12	60.0	4075	3	US-08-796-101-31	Sequence 31, Appl	C 631	12	60.0	9871	3	US-09-479-123-24	Sequence 24, Appl
C 559	12	60.0	4075	3	US-09-085-273-67	Sequence 67, Appl	C 632	12	60.0	9871	3	US-09-484-997-24	Sequence 24, Appl
C 560	12	60.0	4075	3	US-09-916-963-67	Sequence 67, Appl	C 633	12	60.0	9871	3	US-09-481-355-24	Sequence 24, Appl
C 561	12	60.0	4286	3	US-08-227-800A-2	Sequence 2, Appl1	C 634	12	60.0	9871	3	US-09-455-659A-25	Sequence 25, Appl
C 562	12	60.0	4286	3	US-08-921-954-2	Sequence 2, Appl1	C 635	12	60.0	9871	3	US-09-484-996-24	Sequence 24, Appl
C 563	12	60.0	4290	3	US-08-924-629C-4	Sequence 4, Appl1	C 636	12	60.0	9871	3	US-09-484-996-24	Sequence 24, Appl
C 564	12	60.0	4297	3	US-09-710-279-4092	Sequence 4092, Ap	C 637	12	60.0	9871	3	US-09-479-123-24	Sequence 24, Appl
C 565	12	60.0	4459	3	US-09-363-243-2	Sequence 2, Appl1	C 638	12	60.0	9871	3	US-09-484-317A-24	Sequence 24, Appl
C 566	12	60.0	4531	3	US-09-620-312D-893	Sequence 893, Appl	C 639	12	60.0	9871	3	US-09-276-820A-24	Sequence 24, Appl
C 567	12	60.0	4537	3	US-09-254-329-1	Sequence 1, Appl1	C 640	12	60.0	10060	3	US-09-479-122-25	Sequence 25, Appl
C 568	12	60.0	4909	2	US-08-658-665-68	Sequence 68, Appl	C 641	12	60.0	10060	3	US-09-484-997-25	Sequence 25, Appl
C 569	12	60.0	4909	3	US-08-796-101-32	Sequence 32, Appl	C 642	12	60.0	10060	3	US-09-481-355-25	Sequence 25, Appl
C 570	12	60.0	4909	3	US-09-085-273-68	Sequence 68, Appl	C 643	12	60.0	10060	3	US-09-481-355-25	Sequence 25, Appl
C 571	12	60.0	4909	3	US-09-916-963-68	Sequence 68, Appl	C 644	12	60.0	10060	3	US-09-455-659A-25	Sequence 25, Appl
C 572	12	60.0	5111	3	US-09-004-838-118	Sequence 118, Appl	C 645	12	60.0	10060	3	US-09-484-996-25	Sequence 25, Appl
C 573	12	60.0	5443	3	US-09-079-413-5	Sequence 5, Appl1	C 646	12	60.0	10060	3	US-09-479-123-25	Sequence 25, Appl
C 574	12	60.0	5443	3	US-08-750-458A-1	Sequence 1, Appl1	C 647	12	60.0	10060	3	US-09-484-317A-25	Sequence 25, Appl
C 575	12	60.0	6360	3	US-09-171-699-9	Sequence 9, Appl1	C 648	12	60.0	10060	3	US-09-276-820A-25	Sequence 25, Appl
C 576	12	60.0	6436	3	US-09-949-016-16408	Sequence 16408, A	C 649	12	60.0	10580	2	US-08-196-259-1	Sequence 1, Appl1
C 577	12	60.0	6749	2	US-08-658-665-71	Sequence 71, Appl	C 650	12	60.0	10596	2	US-07-885-811-15	Sequence 15, Appl
C 578	12	60.0	6749	2	US-08-658-665-177	Sequence 177, Appl	C 651	12	60.0	10596	2	US-07-885-811-15	Sequence 15, Appl
C 579	12	60.0	6749	3	US-08-796-101-35	Sequence 35, Appl	C 652	12	60.0	10596	2	US-08-087-783A-15	Sequence 15, Appl
C 580	12	60.0	6749	3	US-08-796-101-38	Sequence 38, Appl	C 653	12	60.0	10596	2	US-08-194-088B-15	Sequence 15, Appl
C 581	12	60.0	6749	3	US-09-085-273-71	Sequence 71, Appl	C 654	12	60.0	10596	2	US-08-194-087-15	Sequence 15, Appl
C 582	12	60.0	6749	3	US-09-916-963-71	Sequence 71, Appl	C 655	12	60.0	10596	6	PCT-US93-04648-15	Sequence 15, Appl
C 583	12	60.0	6799	3	US-10-017-754-1883	Sequence 1883, Ap	C 656	12	60.0	11437	3	US-09-949-002-587	Sequence 587, App
C 584	12	60.0	7808	3	US-09-453-702B-247	Sequence 247, Appl	C 657	12	60.0	11563	3	US-09-902-540-1019	Sequence 1019, App
C 585	12	60.0	8374	3	US-10-114-170-247	Sequence 247, Appl	C 658	12	60.0	11589	3	US-09-949-002-826	Sequence 826, App
C 586	12	60.0	8374	3	US-09-949-016-11911	Sequence 11911, A	C 659	12	60.0	11616	2	US-08-196-259-2	Sequence 2, Appl1
C 587	12	60.0	8501	3	US-09-298-367B-6	Sequence 6, Appl1	C 660	12	60.0	12023	3	US-09-949-016-16253	Sequence 16253, A
C 588	12	60.0	8540	3	US-08-487-283A-4	Sequence 4, Appl1	C 661	12	60.0	12270	3	US-09-949-016-16892	Sequence 16892, A
C 589	12	60.0	8540	6	PCT-US96-05611A-12	Sequence 12, Appl	C 662	12	60.0	12270	3	US-09-949-002-830	Sequence 830, App
C 590	12	60.0	8575	6	PCT-US92-08258-6	Sequence 6, Appl1	C 663	12	60.0	12275	3	US-09-949-016-16609	Sequence 16609, A
C 591	12	60.0	8640	3	US-09-964-956-24	Sequence 24, Appl	C 664	12	60.0	12285	3	US-09-949-016-12921	Sequence 12921, A
C 592	12	60.0	8640	3	US-09-964-956-26	Sequence 26, Appl	C 665	12	60.0	12440	3	US-09-949-016-14109	Sequence 14109, A
C 593	12	60.0	8705	3	US-09-647-344A-14	Sequence 14, Appl	C 666	12	60.0	12440	3	US-09-949-016-14110	Sequence 14110, A
C 594	12	60.0	8735	3	US-09-573-080A-19	Sequence 19, Appl	C 667	12	60.0	12440	3	US-09-949-016-14899	Sequence 14899, A
C 595	12	60.0	8831	3	US-09-949-016-15504	Sequence 15504, A	C 668	12	60.0	12440	3	US-09-949-016-14900	Sequence 14900, A
C 596	12	60.0	8885	3	US-09-634-238-26	Sequence 26, Appl	C 669	12	60.0	12759	3	US-09-949-016-16299	Sequence 16299, A
C 597	12	60.0	8932	2	US-08-252-493C-8	Sequence 8, Appl1	C 670	12	60.0	12759	3	US-09-949-016-16300	Sequence 16300, A
C 598	12	60.0	8932	3	US-09-276-197-8	Sequence 8, Appl1	C 671	12	60.0	12759	3	US-09-949-016-16301	Sequence 16301, A
C 599	12	60.0	9132	3	US-09-949-016-13642	Sequence 13642, A	C 672	12	60.0	12759	3	US-09-949-016-16302	Sequence 16302, A
C 600	12	60.0	9132	3	US-09-949-016-13070	Sequence 13070, A	C 673	12	60.0	12759	3	US-09-949-016-16303	Sequence 16303, A
C 601	12	60.0	9600	3	US-08-910-647-1	Sequence 1, Appl1	C 674	12	60.0	12759	3	US-09-949-016-16304	Sequence 16304, A
C 602	12	60.0	9600	3	US-09-620-925-1	Sequence 1, Appl1	C 675	12	60.0	12759	3	US-09-949-016-16305	Sequence 16305, A
C 603	12	60.0	9600	3	US-09-479-122-22	Sequence 22, Appl	C 676	12	60.0	12768	3	US-09-949-016-14302	Sequence 14302, A
C 604	12	60.0	9737	3	US-09-479-122-23	Sequence 23, Appl	C 677	12	60.0	12768	3	US-09-949-016-14303	Sequence 14303, A
C 605	12	60.0	9737	3	US-09-479-122-28	Sequence 28, Appl	C 678	12	60.0	12768	3	US-09-949-016-14304	Sequence 14304, A
C 606	12	60.0	9737	3	US-09-484-997-22	Sequence 22, Appl	C 679	12	60.0	12768	3	US-09-949-016-14305	Sequence 14305, A
C 607	12	60.0	9737	3	US-09-484-997-23	Sequence 23, Appl	C 680	12	60.0	12768	3	US-09-949-016-14306	Sequence 14306, A
C 608	12	60.0	9737	3	US-09-484-997-28	Sequence 28, Appl	C 681	12	60.0	12768	3	US-09-949-016-14307	Sequence 14307, A





c 974 12 60.0 150032 3 US-09-949-016-14321  
c 975 12 60.0 150780 3 US-09-949-016-14711  
976 12 60.0 152393 3 US-09-949-016-14514  
977 12 60.0 152393 3 US-09-949-016-14515  
c 978 12 60.0 154626 3 US-09-949-016-14000  
979 12 60.0 156894 3 US-09-949-016-12765  
980 12 60.0 156894 3 US-09-949-016-12766  
981 12 60.0 156895 3 US-09-949-016-16957  
982 12 60.0 156895 3 US-09-949-016-16958  
983 12 60.0 156895 3 US-09-949-016-16959  
c 984 12 60.0 157866 3 US-09-949-016-12982  
c 985 12 60.0 157866 3 US-09-949-016-12983  
c 986 12 60.0 157866 3 US-09-949-016-12984  
987 12 60.0 161124 3 US-09-949-016-11760  
c 988 12 60.0 161124 3 US-09-949-016-11760  
c 989 12 60.0 162025 3 US-09-834-700-13  
c 990 12 60.0 162025 3 US-09-834-700-14  
c 991 12 60.0 162025 3 US-09-834-700-17  
c 992 12 60.0 162025 3 US-09-834-700-18  
993 12 60.0 162450 3 US-09-345-882-1  
994 12 60.0 162450 3 US-10-071-179-1  
c 995 12 60.0 162841 3 US-09-949-016-13733  
996 12 60.0 163022 3 US-09-949-016-16515  
997 12 60.0 163181 3 US-09-949-016-13730  
c 998 12 60.0 163662 3 US-09-949-016-12545  
c 999 12 60.0 163664 3 US-09-949-016-13546  
c1000 12 60.0 168971 3 US-09-949-016-13807

## ALIGNMENTS

RESULT 1  
US-09-794-928A-12/c  
; Sequence 12, Application US/09794928A  
; Patent No. 6777182  
; GENERAL INFORMATION:  
; APPLICANT: Baban, Soheyl  
; APPLICANT: Bernard, Monique  
; APPLICANT: Cherry, Elana  
; APPLICANT: Gosselin, Diane  
; APPLICANT: Hugo, Patrice  
; APPLICANT: Malette, Brigitte  
; APPLICANT: Miron, Pierre  
; APPLICANT: Prive, Charles  
; APPLICANT: Shazand, Kamran  
; TITLE OF INVENTION: ENDOMETRIOSIS-RELATED MARKERS AND USES THEREOF  
; FILE REFERENCE: 5600.71  
; CURRENT APPLICATION NUMBER: US/09/794,928A  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: US 60/225,745  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: US 60/185,063  
; PRIOR FILING DATE: 2000-02-25  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 254  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-794-928A-12

Query Match 100.0%; Score 20; DB 3; Length 254;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20  
|||||  
Db 90 GCTTTGGTTGGGCAACACAT 71

RESULT 2  
US-09-949-016-16598

; Sequence 16598, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16598  
; LENGTH: 10603  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16598  
  
Query Match 80.0%; Score 16; DB 3; Length 10603;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TTGGTTGGGCAACACA 19  
|||||  
Db 6431 TTGGTTGGGCAACACA 6446

## RESULT 3

US-09-949-016-42568  
; Sequence 42568, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42568  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-42568

Query Match 75.0%; Score 15; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TTGGTTGGGCAACACA 19  
|||||  
Db 77 TTGGTTGGGCAACACA 91

## RESULT 4

US-08-976-259-66  
; Sequence 66, Application US/08976259  
; Patent No. 6316609  
; GENERAL INFORMATION:  
; APPLICANT: Dillon, Patrick J.  
; APPLICANT: Choi, Gil H.

APPLICANT: Welch, Rodney A.  
TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli  
Patent No. 6316609  
NUMBER OF SEQUENCES: 142  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Ave, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
FILING DATE: Herewith  
APPLICATION NUMBER: US/08/976,259  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/031,626 AND US 60/061,953  
ATTORNEY/AGENT INFORMATION:  
NAME: Steffe, Eric K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.0740002/EKS/CEM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1342 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-976-259-66

Query Match 75.0%; Score 15; DB 3; Length 1342;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAA 15  
|||||  
Db 614 GCTTTGGTTGGCAA 628

RESULT 5  
US-09-956-004-66  
Sequence 66, Application US/09956004  
Patent No. 6787643  
GENERAL INFORMATION:  
APPLICANT: Patrick J. Dillon et al.  
TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands  
FILE REFERENCE: PB324D1  
CURRENT APPLICATION NUMBER: US/09/956,004  
CURRENT FILING DATE: 2001-09-20  
PRIOR APPLICATION NUMBER: 08/976,259  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/061,953  
PRIOR FILING DATE: 1997-10-14  
PRIOR APPLICATION NUMBER: 60/031,626  
PRIOR FILING DATE: 1996-11-22  
NUMBER OF SEQ ID NOS: 142  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 66  
LENGTH: 1342  
TYPE: DNA  
ORGANISM: Escherichia coli  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1238)..(1238)  
OTHER INFORMATION: n equals a, t, g, or c  
US-09-956-004-66

Query Match 75.0%; Score 15; DB 3; Length 1342;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAA 15  
|||||  
Db 614 GCTTTGGTTGGCAA 628

RESULT 6  
US-09-543-681A-3630  
Sequence 3630, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 3630  
LENGTH: 1701  
TYPE: DNA  
ORGANISM: Proteus mirabilis  
US-09-543-681A-3630

Query Match 75.0%; Score 15; DB 3; Length 1701;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TTGGTTGGCAACAC 18  
|||||  
Db 1271 TTGGTTGGCAACAC 1285

RESULT 7  
US-09-328-352-3989  
Sequence 3989, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 3989  
LENGTH: 2181  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-3989

Query Match 75.0%; Score 15; DB 3; Length 2181;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAA 15  
|||||  
Db 884 GCTTTGGTTGGCAA 898

RESULT 8  
US-09-645-069-22  
Sequence 22, Application US/09645069  
Patent No. 6808710  
GENERAL INFORMATION:  
APPLICANT: Clive Wood  
APPLICANT: Gordon Freeman

; TITLE OF INVENTION: PD-1, A Receptor For B7-4, and Uses Therefor  
; FILE REFERENCE: GNN-004B  
; CURRENT APPLICATION NUMBER: US/09/645,069  
; CURRENT FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: 60/150,390  
; PRIOR FILING DATE: 1999-8-23  
; PRIOR APPLICATION NUMBER: 60/164,897  
; PRIOR FILING DATE: 1999-11-10  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 3593  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (17)..(889)  
US-09-645-069-22

Query Match 75.0%; Score 15; DB 3; Length 3593;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 3406 GGTGGGCAACACAT 3420

RESULT 9  
US-09-644-934-10  
; Sequence 10, Application US/09644934  
; Patent No. 6936704  
; GENERAL INFORMATION:  
; APPLICANT: Gordon Freeman  
; APPLICANT: Vassiliki Boussiotis  
; APPLICANT: Tatyana Chernova  
; APPLICANT: Nelly Malenkovich  
; TITLE OF INVENTION: NOVEL B7-4 FAMILY MOLECULES AND USES THEREFOR  
; FILE REFERENCE: GNN-004A  
; CURRENT APPLICATION NUMBER: US/09/644,934  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/150,390  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 3593  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (17)..(889)  
US-09-644-934-10

Query Match 75.0%; Score 15; DB 3; Length 3593;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 3406 GGTGGGCAACACAT 3420

RESULT 10  
US-09-724-126A-18  
; Sequence 18, Application US/09724126A  
; Patent No. 6706505  
; GENERAL INFORMATION:  
; APPLICANT: Han, Hui-Quan  
; APPLICANT: Kwak, Keith  
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family  
; FILE REFERENCE: 01017/35966A  
; CURRENT APPLICATION NUMBER: US/09/724,126A

; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,211  
; PRIOR FILING DATE: 1999-03-01  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 18  
; LENGTH: 5205  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (662)  
; OTHER INFORMATION: n = A or T or G or C  
; NAME/KEY: misc feature  
; LOCATION: (668)  
; OTHER INFORMATION: n = A or T or G or C  
US-09-724-126A-18

Query Match 75.0%; Score 15; DB 3; Length 5205;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1518 GGTGGGCAACACAT 1532

RESULT 11  
US-09-724-126A-1  
; Sequence 1, Application US/09724126A  
; Patent No. 6706505  
; GENERAL INFORMATION:  
; APPLICANT: Han, Hui-Quan  
; APPLICANT: Kwak, Keith  
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family  
; FILE REFERENCE: 01017/35966A  
; CURRENT APPLICATION NUMBER: US/09/724,126A  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,211  
; PRIOR FILING DATE: 1999-03-01  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 6308  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (696)..(5942)  
US-09-724-126A-1

Query Match 75.0%; Score 15; DB 3; Length 6308;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 2258 GGTGGGCAACACAT 2272

RESULT 12  
US-09-949-016-12968/c  
; Sequence 12968, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20



; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12968  
; LENGTH: 222452  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12968

Query Match 75.0%; Score 15; DB 3; Length 222452;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGCAACACA 19  
|||||

Db 179856 TGGTTGGCAACACA 179842

## RESULT 13

US-09-557-884-1  
; Sequence 1, Application US/09557884  
; Patent No. 6506581  
; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: The Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
Thereof, and Uses Thereof

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: MD

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3 1/2 inch diskette

; COMPUTER: Dell Pentium

; OPERATING SYSTEM: MS DOS v6.22

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/557,884

; FILING DATE: 25-Apr-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/476,102

; FILING DATE: JUN-5-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

; REFERENCE/DOCKET NUMBER: PB186P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 301-309-8504

; TELEFAX: 301-309-8439

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1830121 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 75.0%; Score 15; DB 3; Length 1830121;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAA 15  
|||||

Db 1130570 GCTTTGGTTGGGCAA 1130584

## RESULT 14

US-09-643-990A-1

; Sequence 1, Application US/09643990A

; Patent No. 6528289

; GENERAL INFORMATION:

; APPLICANT: Robert D. Fleischmann

; Mark D. Adams

; Owen White

; Hamilton O. Smith

; J. Craig Venter

; TITLE OF INVENTION: The Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
Thereof, and Uses Thereof

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville,

; STATE: MD

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3 1/2 inch diskette

; COMPUTER: Dell Pentium

; OPERATING SYSTEM: MS DOS v6.22

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/643,990A

; FILING DATE: 23-Aug-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/487,429

; FILING DATE: 1995-06-07

; APPLICATION NUMBER: 08/426,787

; FILING DATE: 1995-04-21

; ATTORNEY/AGENT INFORMATION:

; NAME: Kenley K. Hoover

; REGISTRATION NUMBER: 40,302

; REFERENCE/DOCKET NUMBER: PB186P1C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 301-610-5790

; TELEFAX: 310-309-8439

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1830121 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Query Match 75.0%; Score 15; DB 3; Length 1830121;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAA 15  
|||||

Db 1130570 GCTTTGGTTGGGCAA 1130584

## RESULT 15

US-10-158-865-1

; Sequence 1, Application US/10158865

; Patent No. 6846651

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag  
ments Thereof, and Uses Thereof

; Patent No. 6846651

; TITLE OF INVENTION: Thereof, and Uses Thereof

; FILE REFERENCE: PB186P2C1D1

; CURRENT APPLICATION NUMBER: US/10/158,865

; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US 09/557,884  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: US 08/476,102  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: US 08/426,787  
; PRIOR FILING DATE: 1995-04-21  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1830121  
; TYPE: DNA  
; ORGANISM: Haemophilus influenzae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4747)..(4747)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (9921)..(9921)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (10150)..(10150)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (29298)..(29298)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36543)..(36543)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36551)..(36551)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36636)..(36636)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (40808)..(40810)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (44416)..(44416)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (44905)..(44905)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (44975)..(44975)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (45593)..(45593)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (45732)..(45732)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (47036)..(47036)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (51334)..(51334)  
; OTHER INFORMATION: n equals a,t,c, or g

; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (51602)..(51602)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (51786)..(51786)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (51805)..(51805)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (55369)..(55369)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (65309)..(65309)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (65313)..(65313)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (80024)..(80024)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (100091)..(100091)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (102696)..(102696)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (105121)..(105121)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (107248)..(107248)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (117136)..(117136)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (119750)..(119750)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (119924)..(119924)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (120038)..(120038)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (121344)..(121344)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (122167)..(122167)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (122336)..(122336)  
; OTHER INFORMATION: n equals a,t,c, or g  
; FEATURE:



```
RESULT 18
US-09-188-930-77
; Sequence 77, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188.930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 77
; LENGTH: 440
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (9)...(9)
US-09-188-930-77

Query Match          70.0%; Score 14; DB 3; Length 440;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 TTTGTTGGGCAAC 16
Db      401 TTGTTGGGCAAC 414

RESULT 19
US-09-949-016-27657/c
; Sequence 27657, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27657
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-27657

Query Match          70.0%; Score 14; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 GGTGGGCAACACA 19
Db      284 GGTGGGCAACACA 271

RESULT 20
US-09-949-016-34062
; Sequence 34062, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34062
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34062

Query Match          70.0%; Score 14; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      5 TGGTTGGGCAACAC 18
Db      413 TGGTTGGGCAACAC 426

RESULT 21
US-09-949-016-158206/c
; Sequence 158206, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 158206
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-158206

Query Match          70.0%; Score 14; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 GGTGGGCAACACA 19
Db      284 GGTGGGCAACACA 271

RESULT 22
US-09-949-016-186133
; Sequence 186133, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 186133
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-186133

Query Match          70.0%; Score 14; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TGGTTGGGCAAC 18
Db 413 TGGTTGGGCAAC 426

RESULT 23
US-09-442-143A-7
; Sequence 7, Application US/09442143A
; Patent No. 6403089
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-14
; CURRENT APPLICATION NUMBER: US/09/442,143A
; CURRENT FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Murine
US-09-442-143A-7

Query Match          70.0%; Score 14; DB 3; Length 707;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TTTGGTTGGGCAAC 16
Db 232 TTTGGTTGGGCAAC 245

RESULT 24
US-09-902-563-7
; Sequence 7, Application US/09902563
; Patent No. 6805863
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-37
; CURRENT APPLICATION NUMBER: US/09/902,563
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: US 09/442,143
; PRIOR FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 7
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Murine
US-09-902-563-7

Query Match          70.0%; Score 14; DB 3; Length 707;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TTTGGTTGGGCAAC 16
Db 232 TTTGGTTGGGCAAC 245

RESULT 25
US-09-583-110-2269
; Sequence 2269, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2269
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2269

Query Match          70.0%; Score 14; DB 3; Length 1194;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TTTGGTTGGGCAAC 16
Db 448 TTTGGTTGGGCAAC 461

RESULT 26
US-09-107-433-294
; Sequence 294, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 294:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1194
; SEQUENCE DESCRIPTION: SEQ ID NO: 294:
US-09-107-433-294

Query Match 70.0%; Score 14; DB 3; Length 1194;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTTGGTTGGGCAAC 16
Db 448 TTTGGTTGGGCAAC 461

RESULT 27
US-08-961-083-47/c
; Sequence 47, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE: OCT-30-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-536-784-47

Query Match 70.0%; Score 14; DB 3; Length 1279;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTTGGTTGGGCAAC 16
Db 273 TTTGGTTGGGCAAC 260

RESULT 29
US-09-765-271-47/c
; Sequence 47, Application US/09765271
; Patent No. 6887663
; GENERAL INFORMATION:
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-083-47

Query Match 70.0%; Score 14; DB 3; Length 1279;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTTGGTTGGGCAAC 16
Db 273 TTTGGTTGGGCAAC 260

RESULT 28
US-09-536-784-47/c
; Sequence 47, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-536-784-47

Query Match 70.0%; Score 14; DB 3; Length 1279;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTTGGTTGGGCAAC 16
Db 273 TTTGGTTGGGCAAC 260

RESULT 29
US-09-765-271-47/c
; Sequence 47, Application US/09765271
; Patent No. 6887663
; GENERAL INFORMATION:
```

```
;
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,271
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/536,784
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
;
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-765-271-47

      Query Match          70.0%; Score 14; DB 3; Length 1279;
      Best Local Similarity 100.0%; Pred. No. 47;
      Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 TTTGGTTGGCAAC 16
Db      273 TTTGGTTGGCAAC 260

RESULT 30
US-09-765-272A-47/c
; Sequence 47, Application US/09765272A
; Patent No. 6929930
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 454
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: Dell Latitude C610
; OPERATING SYSTEM: Windows 2000
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272A
; FILING DATE: 22-Jan-2001
```

```
;
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin J. Hymel
; REGISTRATION NUMBER: 45,414
; REFERENCE/DOCKET NUMBER: PB340P2C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 610-5790
; TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-765-272A-47

      Query Match          70.0%; Score 14; DB 3; Length 1279;
      Best Local Similarity 100.0%; Pred. No. 47;
      Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 TTTGGTTGGCAAC 16
Db      273 TTTGGTTGGCAAC 260

RESULT 31
US-09-196-857-1/c
; Sequence 1, Application US/09196857A
; Patent No. 6100069
; GENERAL INFORMATION:
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Brown, James R.
; APPLICANT: Biswas, Sanjoy
; APPLICANT: Throup, John P.
; APPLICANT: Lawlor, Elizabeth J.
; APPLICANT: Mooney, Jeffrey
; APPLICANT: Zhong, Yiyi
; APPLICANT: Debouck, Christine
; APPLICANT: Jaworski, Deborah D.
; APPLICANT: Wang, Min
; APPLICANT: Warren, Richard L.
; APPLICANT: Schilling, Lisa K.
; TITLE OF INVENTION: No. 6100069el tig
; FILE REFERENCE: GM10118
; CURRENT APPLICATION NUMBER: US/09/196,857A
; CURRENT FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/082,418
; EARLIER FILING DATE: 1997-11-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1281)
;
; US-09-196-857-1

      Query Match          70.0%; Score 14; DB 3; Length 1284;
      Best Local Similarity 100.0%; Pred. No. 47;
      Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 TTTGGTTGGCAAC 16
Db      275 TTTGGTTGGCAAC 262

RESULT 32
```

```
US-09-583-110-1279/c
; Sequence 1279, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1279
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-1279

Query Match          70.0%; Score 14; DB 3; Length 1284;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 TTTGGTTGGGCAAC 16
      |||||
Db      275 TTTGGTTGGGCAAC 262

RESULT 33
US-09-107-433-1653/c
; Sequence 1653, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1653:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1308 base pairs
; TYPE: nucleic acid
; . . .
```

```
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1308
; SEQUENCE DESCRIPTION: SEQ ID NO: 1653:
US-09-107-433-1653

Query Match          70.0%; Score 14; DB 3; Length 1308;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 TTTGGTTGGGCAAC 16
      |||||
Db      299 TTTGGTTGGGCAAC 286

RESULT 34
US-09-917-254-22/c
; Sequence 22, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1449
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (1316)..(1316)
; OTHER INFORMATION: n = a, c, g, or t
; NAME/KEY: Unsure
; LOCATION: (1360)..(1360)
; OTHER INFORMATION: n = a, c, g, or t
; NAME/KEY: Unsure
; LOCATION: (1366)..(1367)
; OTHER INFORMATION: n = a, c, g, or t
; NAME/KEY: Unsure
; LOCATION: (1369)..(1369)
; OTHER INFORMATION: n = a, c, g, or t
US-09-917-254-22

Query Match          70.0%; Score 14; DB 3; Length 1449;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 TTTGGTTGGGCAAC 16
      |||||
Db      1429 TTTGGTTGGGCAAC 1416

RESULT 35
US-09-949-016-4402/c
; Sequence 4402, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```



FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4402  
LENGTH: 2184  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-4402

Query Match 70.0%; Score 14; DB 3; Length 2184;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TGGTTGGGCAACAC 18  
|||||  
DB 1094 TGGTTGGGCAACAC 1081

RESULT 36  
US-08-396-479B-3  
Sequence 3, Application US/08396479B  
Patent No. 5612455  
GENERAL INFORMATION:  
APPLICANT: HOEY, Timothy  
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/396,479B  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59450-1/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 210 277299  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2743 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 240..2390  
US-08-396-479B-3

Query Match 70.0%; Score 14; DB 2; Length 2743;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TGGTTGGGCAACAC 18  
|||||  
DB 1137 TGGTTGGGCAACAC 1150

RESULT 37  
US-08-818-823-3  
Sequence 3, Application US/08818823  
Patent No. 5708158  
GENERAL INFORMATION:  
APPLICANT: HOEY, Timothy  
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,823  
FILING DATE: 14-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/396,479  
FILING DATE: 02-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59450-1/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 210 277299  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2743 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 240..2390  
US-08-818-823-3

Query Match 70.0%; Score 14; DB 2; Length 2743;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TGGTTGGGCAACAC 18  
|||||  
DB 1137 TGGTTGGGCAACAC 1150

RESULT 38  
US-08-124-981A-1  
Sequence 1, Application US/08124981A  
Patent No. 5837840  
GENERAL INFORMATION:  
APPLICANT: Crabtree, Gerald R.  
APPLICANT: No. 5837840thorp, Jeffrey P.  
TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew

STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/124,981A  
FILING DATE: 20-SEP-1993  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 5490A-226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2749 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 103..2385  
US-08-124-981A-1

Query Match 70.0%; Score 14; DB 2; Length 2749;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGGCAACAC 18  
|||||  
Db 1135 TGGTTGGGCAACAC 1148

RESULT 39  
US-09-037-190-45  
Sequence 45, Application US/09037190  
Patent No. 6096515  
GENERAL INFORMATION:  
APPLICANT: Crabtree, Gerald R.  
APPLICANT: No. 6096515throp, Jeffrey P.  
APPLICANT: Ho, Steffan M.  
TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/037,190  
FILING DATE: 09-MAR-1998  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,174  
FILING DATE: 13-JUN-1994  
APPLICATION NUMBER: US 08/124,981

FILING DATE: 20-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: APV-332.03  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2751 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 240..2387  
US-09-037-190-45

Query Match 70.0%; Score 14; DB 3; Length 2751;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGGCAACAC 18  
|||||  
Db 1137 TGGTTGGGCAACAC 1150

RESULT 40  
US-09-037-192-45  
Sequence 45, Application US/09037192  
Patent No. 6096860  
GENERAL INFORMATION:  
APPLICANT: Crabtree, Gerald R.  
APPLICANT: No. 6096860throp, Jeffrey P.  
APPLICANT: Ho, Steffan M.  
TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/037,192  
FILING DATE: 09-MAR-1998  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,174  
FILING DATE: 13-JUN-1994  
APPLICATION NUMBER: US 08/124,981  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: APV-332.04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2751 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 240..2387
US-09-037-192-45

Query Match          70.0%; Score 14; DB 3; Length 2751;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGGCAACAC 18
    |||||
Db 1137 TGGTTGGGCAACAC 1150

RESULT 41
US-09-037-143-45
; Sequence 45, Application US/09037143A
; Patent No. 6150099
; GENERAL INFORMATION:
; APPLICANT: Crabtree, Gerald R.
; APPLICANT: No. 6150099throp, Jeffrey P.
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: APV-332.05
; CURRENT APPLICATION NUMBER: US/09/037,143A
; EARLIER APPLICATION NUMBER:
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 2751
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (240)..(2387)
US-09-037-143-45

Query Match          70.0%; Score 14; DB 3; Length 2751;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGGCAACAC 18
    |||||
Db 1137 TGGTTGGGCAACAC 1150

RESULT 42
US-09-049-691-45
; Sequence 45, Application US/09049691
; Patent No. 6171781
; GENERAL INFORMATION:
; APPLICANT: Crabtree, Gerald R.
; APPLICANT: No. 617178throp, Jeffrey P.
; APPLICANT: Ho, Steffan M.
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HONG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/049,691
; FILING DATE: 27-MAR-1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,174
; FILING DATE: 13-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/124,981
; FILING DATE: 20-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: APV-332.09
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2751 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 240..2387
US-09-049-691-45

Query Match          70.0%; Score 14; DB 3; Length 2751;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGGCAACAC 18
    |||||
Db 1137 TGGTTGGGCAACAC 1150

RESULT 43
US-08-260-174-45
; Sequence 45, Application US/08260174C
; Patent No. 6197925
; GENERAL INFORMATION:
; APPLICANT: Crabtree, Gerald R.
; APPLICANT: No. 6197925throp, Jeffrey P.
; APPLICANT: Ho, Steffan M.
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: APV-332.02
; CURRENT APPLICATION NUMBER: US/08/260,174C
; CURRENT FILING DATE: 1994-06-13
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 2751
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (240)..(2387)
US-08-260-174-45

Query Match          70.0%; Score 14; DB 3; Length 2751;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGGCAACAC 18
    |||||
Db 1137 TGGTTGGGCAACAC 1150

RESULT 44
US-09-338-128A-45
```

; Sequence 45, Application US/09338128A  
; Patent No. 6312899  
; GENERAL INFORMATION:  
; APPLICANT: Crabtree, Gerald R.  
; APPLICANT: No. 6312899throp, Jeffrey P.  
; APPLICANT: Ho, Steffan M.  
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES  
; FILE REFERENCE: APV-33211  
; CURRENT APPLICATION NUMBER: US/09/338,128A  
; CURRENT FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 08/260,174  
; PRIOR FILING DATE: 1994-06-13  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 2751  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (240)..(2387)  
US-09-338-128A-45

Query Match 70.0%; Score 14; DB 3; Length 2751;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGGCAACAC 18  
Db 1137 TGGTTGGGCAACAC 1150

RESULT 45  
US-09-037-192-45  
; Sequence 45, Application US/09037192  
; Patent No. 6388052  
; GENERAL INFORMATION:  
; APPLICANT: Crabtree, Gerald R.  
; Ho, Steffan M.  
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/037,192  
; FILING DATE: 09-Mar-1998  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/260,174  
; FILING DATE: 13-JUN-1994  
; APPLICATION NUMBER: US 08/124,981  
; FILING DATE: 20-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: APV-332.04  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2751 base pairs

; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 240..2387  
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:  
US-09-037-192-45

Query Match 70.0%; Score 14; DB 3; Length 2751;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGGCAACAC 18  
Db 1137 TGGTTGGGCAACAC 1150

RESULT 46  
PCT-US94-07297-36  
; Sequence 36, Application PC/TUS9407297  
; GENERAL INFORMATION:  
; APPLICANT: Arai, Naoko  
; APPLICANT: Masuda, Eateban S.  
; APPLICANT: Tokumitsu, Hiroshi  
; TITLE OF INVENTION: PURIFIED COMPONENTS OF MAMMALIAN  
; TITLE OF INVENTION: TRANSCRIPTION REGULATION COMPLEXES, AND ANALOGS  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John H. C. Blasdale, Schering-Plough Corporation, M-3-W  
; STREET: One Giralda Farms  
; CITY: Madison  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07940-1000  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh IICx  
; OPERATING SYSTEM: System Software 7.1  
; SOFTWARE: Microsoft Word 5.1a  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/07297  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/222,626  
; FILING DATE: 04-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,061  
; FILING DATE: 05-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/113,971  
; FILING DATE: 30-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/099,998  
; FILING DATE: 30-JUL-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/088,483  
; FILING DATE: 06-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Blasdale, John H. C.  
; REGISTRATION NUMBER: 31,895  
; REFERENCE/DOCKET NUMBER: DX0392K4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-822-7398  
; TELEFAX: 201-822-7039  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2853 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

```

; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 340...2490
PCT-US94-07297-36

Query Match          70.0%; Score 14; DB 6; Length 2853;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TGGTTGGGCAACAC 18
Db 1237 TGGTTGGGCAACAC 1250

RESULT 47
US-09-949-016-823
; Sequence 823, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 823
; LENGTH: 2935
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-823

Query Match          70.0%; Score 14; DB 3; Length 2935;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TGGTTGGGCAACAC 18
Db 934 TGGTTGGGCAACAC 947

RESULT 48
US-09-220-132-68
; Sequence 68, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyian, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/069,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 3082
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-68

Query Match          70.0%; Score 14; DB 3; Length 3082;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TGGTTGGGCAACAC 18
Db 1091 TGGTTGGGCAACAC 1104

Search completed: May 6, 2006, 17:19:57
Job time : 129 secs
```

```

Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GTTGGGCAACACAT 20
Db 406 GTTGGGCAACACAT 419

RESULT 49
US-09-540-236-1660
; Sequence 1660, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1660
; LENGTH: 4152
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-1660

Query Match          70.0%; Score 14; DB 3; Length 4152;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TTTGGTTGGGCAAC 16
Db 3438 TTTGGTTGGGCAAC 3451

RESULT 50
US-09-949-016-5322
; Sequence 5322, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5322
; LENGTH: 4464
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5322

Query Match          70.0%; Score 14; DB 3; Length 4464;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TGGTTGGGCAACAC 18
Db 1091 TGGTTGGGCAACAC 1104

Search completed: May 6, 2006, 17:19:57
Job time : 129 secs
```

THIS PAGE LEFT BLANK

GenCore version 5.1.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2006, 00:46:02 ; Search time 1768 Seconds  
(without alignments)  
643.025 Million cell updates/sec

Title: US-10-671-074-172

Perfect score: 20

Sequence: 1 gcttgggtggcaacacat 20

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 1

Total number of hits satisfying chosen parameters: 11765110

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.in.\*
- 3: gb.env.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pr.\*
- 9: gb.pro.\*
- 10: gb.sts.\*
- 11: gb.sy.\*
- 12: gb.un.\*
- 13: gb.vi.\*
- 14: gb.htg.\*
- 15: gb.pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	254	6 AR576665	Sequence
C 2	20	100.0	254	6 AX229886	Sequence
C 3	20	100.0	303	6 C0679217	Sequence
C 4	20	100.0	447	6 C0690723	Sequence
C 5	20	100.0	478	6 AX778044	Sequence
C 6	20	100.0	1968	8 BT007455	Homo sapi
C 7	20	100.0	1968	11 AY890173	Synthetic
C 8	20	100.0	1968	11 AY890174	Synthetic
C 9	20	100.0	2103	9 AF126056	Mus muscu
C 10	20	100.0	2162	4 AY094061	Sus scro
C 11	20	100.0	2413	8 BC021981	Homo sapi
C 12	20	100.0	2460	9 AF114258	Mus muscu
C 13	20	100.0	2482	8 BC070065	Homo sapi
C 14	20	100.0	2503	9 AY255525	Spermophi
C 15	20	100.0	2725	11 AF178854	Synthetic
C 16	20	100.0	3200	8 HSU02368	Human PAX3/
C 17	20	100.0	3421	6 C0874011	Sequence
C 18	20	100.0	3421	8 HSU02310	Human fork

C 19	20	100.0	3517	8 HSU02308	Human PAX-3
C 20	20	100.0	4714	6 CQ724998	Sequence
C 21	20	100.0	4919	6 HSM808426	Homo sapi
C 22	20	100.0	4945	9 NMU252157	Mus muscu
C 23	20	100.0	5723	6 CS031181	Sequence
C 24	20	100.0	5723	6 CS040133	Sequence
C 25	20	100.0	5723	6 AX587664	Sequence
C 26	20	100.0	5723	8 AF032885	Homo sapi
C 27	20	100.0	5769	6 AX779990	Sequence
C 28	20	100.0	81862	8 AL133318	Human DNA
C 29	20	100.0	149841	14 AC163092	Mus muscu
C 30	20	100.0	155595	14 AC078964	Homo sapi
C 31	20	100.0	186044	14 AC084099	Homo sapi
C 32	20	100.0	228076	14 AC089851	Rattus no
C 33	20	100.0	303046	14 AC113811	Continuation (12 o
C 34	19	95.0	110000	1 AE016828	Continuation (10 o
C 35	17	85.0	299	15 AJ834861	Arabidops
C 36	17	85.0	2000	6 AX509544	Sequence
C 37	17	85.0	7618	5 AF537216	Sequence
C 38	17	85.0	40406	5 FRIGVH1	Takifugu
C 39	17	85.0	78423	15 AB024035	Pugu rubr
C 40	17	85.0	110000	1 CP000084	Arabidops
C 41	16	80.0	47740	1 AY144116	Continuation (9 of
C 42	16	80.0	105083	14 AC095920	Photorhab
C 43	16	80.0	110000	15 AP008214	Continuation (27 o
C 44	16	80.0	139455	8 AC097534	Continuation (27 o
C 45	16	80.0	141257	15 AP005741	Homo sapi
C 46	16	80.0	154612	14 AC093739	Oryza sat
C 47	16	80.0	160625	8 AC009311	AC093739 Homo sapi
C 48	16	80.0	166472	14 AC009588	AC009311 Homo sapi
C 49	16	80.0	188776	9 BX537258	AC009588 Homo sapi
C 50	16	80.0	193506	9 AC116953	Mouse DNA
C 51	16	80.0	221363	5 BX005389	Mus muscu
C 52	16	80.0	239627	14 AC152177	Zebracfish
C 53	16	80.0	250128	14 AC162246	AC152177 Medicago
C 54	16	80.0	255453	14 AC109391	AC162246 Bos tauru
C 55	16	80.0	261001	14 AC096324	AC109391 Rattus no
C 56	16	80.0	280615	14 AC096297	AC096324 Rattus no
C 57	15	75.0	406	1 AF536051	AC096297 Rattus no
C 58	15	75.0	406	1 AF536052	AF536051 Haemophil
C 59	15	75.0	406	1 AF536053	AF536052 Haemophil
C 60	15	75.0	406	1 AF536054	AF536053 Haemophil
C 61	15	75.0	406	1 AF536055	AF536054 Haemophil
C 62	15	75.0	406	1 AF536056	AF536055 Haemophil
C 63	15	75.0	406	1 AF536057	AF536056 Haemophil
C 64	15	75.0	406	1 AF536058	AF536057 Haemophil
C 65	15	75.0	406	1 AF536059	AF536058 Haemophil
C 66	15	75.0	406	1 AF536060	AF536059 Haemophil
C 67	15	75.0	406	1 AF536061	AF536060 Haemophil
C 68	15	75.0	406	1 AF536062	AF536061 Haemophil
C 69	15	75.0	406	1 AF536063	AF536062 Haemophil
C 70	15	75.0	406	1 AF536064	AF536063 Haemophil
C 71	15	75.0	406	1 AF536065	AF536064 Haemophil
C 72	15	75.0	406	1 AF536066	AF536065 Haemophil
C 73	15	75.0	406	1 AF536067	AF536066 Haemophil
C 74	15	75.0	406	1 AF536068	AF536067 Haemophil
C 75	15	75.0	406	1 AF536069	AF536068 Haemophil
C 76	15	75.0	406	1 AF536070	AF536069 Haemophil
C 77	15	75.0	579	10 G97533	AF536070 Haemophil
C 78	15	75.0	627	10 BV412283	G97533 S208P6567FG
C 79	15	75.0	703	8 HSA328477	BV412283 S229P668F
C 80	15	75.0	714	15 AF482832	AJ328477 Homo sapi
C 81	15	75.0	756	10 BV624260	AF482832 Agarticus
C 82	15	75.0	817	10 BV665156	BV624260 S215P6161
C 83	15	75.0	1342	6 BD195279	BV665156 S216P6206
C 84	15	75.0	1342	6 BD195279	BD195279 Nucleotid
C 85	15	75.0	1413	6 CQ739385	BD195279 Sequence
C 86	15	75.0	1513	8 BC062359	CQ739385 Sequence
C 87	15	75.0	1550	6 AX834422	BC062359 Homo sapi
C 88	15	75.0	1550	6 AX834422	D89249 Schistosacch
C 89	15	75.0	1550	6 AK096872	AX834422 Sequence
C 90	15	75.0	1701	6 AK378624	AK96872 Homo sapi
C 91	15	75.0	1758	9 BC034067	AK378624 Sequence

c 92	15	75.0	1758	9	BC040758	BC040758 Mus muscu	c 165	15	75.0	166007	8	HSBA425A6	AL121749 Human DNA
c 93	15	75.0	2181	6	AR321439	AR321439 Sequence	c 166	15	75.0	166518	8	AC090514	AC090514 Homo sapi
c 94	15	75.0	2550	6	BD156870	BD156870 Primer fo	c 167	15	75.0	168362	8	AC119228	AC119228 Mus muscu
c 95	15	75.0	2550	6	AX877826	AX877826 Sequence	c 168	15	75.0	168587	8	AP003439	AP003439 Homo sapi
c 96	15	75.0	2550	6	AX027803	AX027803 Homo sapi	c 169	15	75.0	169077	8	AC066613	AC066613 Homo sapi
c 97	15	75.0	3059	6	AX714232	AX714232 Sequence	c 170	15	75.0	170233	15	AC135418	AC135418 Oryza sat
c 98	15	75.0	3059	6	AX056441	AX056441 Homo sapi	c 171	15	75.0	171113	14	AC034266	AC034266 Homo sapi
c 99	15	75.0	3342	6	C0604796	C0604796 Sequence	c 172	15	75.0	171962	14	AC013537	AC013537 Homo sapi
c 100	15	75.0	3366	6	C0595320	C0595320 Sequence	c 173	15	75.0	173814	14	AC093840	AC093840 Homo sapi
c 101	15	75.0	3416	2	AF181657	AF181657 Drosophil	c 174	15	75.0	173476	2	AC092230	AC092230 Drosophil
c 102	15	75.0	3593	6	C0777454	C0777454 Sequence	c 175	15	75.0	175461	14	AC163768	AC163768 Pan trogl
c 103	15	75.0	3593	6	AX592428	AX592428 Sequence	c 176	15	75.0	175488	14	AC013556	AC013556 Homo sapi
c 104	15	75.0	3593	6	AX088418	AX088418 Sequence	c 177	15	75.0	177187	14	AC096848	AC096848 Papio anu
c 105	15	75.0	3593	6	AX088429	AX088429 Sequence	c 178	15	75.0	177242	14	AC021923	AC021923 Homo sapi
c 106	15	75.0	3593	9	AF233517	AF233517 Mus muscu	c 179	15	75.0	178348	14	AC132931	AC132931 Mus muscu
c 107	15	75.0	3697	9	BC066841	BC066841 Mus muscu	c 180	15	75.0	179060	14	AC074382	AC074382 Homo sapi
c 108	15	75.0	4529	6	C0728565	C0728565 Sequence	c 181	15	75.0	179080	14	AC092514	AC092514 Papio anu
c 109	15	75.0	5128	8	AF525401	AF525401 Homo sapi	c 182	15	75.0	179485	8	AC005939	AC005939 Homo sapi
c 110	15	75.0	5205	6	AX487565	AX487565 Sequence	c 183	15	75.0	183773	9	AC154213	AC154213 Mus muscu
c 111	15	75.0	5250	8	AY061886	AY061886 Homo sapi	c 184	15	75.0	184650	14	AC161833	AC161833 Bos tauru
c 112	15	75.0	5492	6	C0595319	C0595319 Sequence	c 185	15	75.0	185137	5	AL954187	AL954187 Zebrafish
c 113	15	75.0	6308	6	AF487554	AF487554 Sequence	c 186	15	75.0	185583	9	AC153553	AC153553 Mus muscu
c 114	15	75.0	12000	5	XLA315159	AX315159 Xenopus 1	c 187	15	75.0	187349	8	AP003352	AP003352 Homo sapi
c 115	15	75.0	12361	1	U32787	U32787 Haemophilus	c 188	15	75.0	188177	9	AC121135	AC121135 Mus muscu
c 116	15	75.0	22469	2	U80033	U80033 Caenorhabdi	c 189	15	75.0	188764	14	AC021238	AC021238 Homo sapi
c 117	15	75.0	23733	15	MAPESYNT	X89442 Metarhizium	c 190	15	75.0	192649	9	AC161805	AC161805 Mus muscu
c 118	15	75.0	47079	14	AC083776	AC083776 Homo sapi	c 191	15	75.0	193326	14	AC127443	AC127443 Rattus no
c 119	15	75.0	53881	14	AP007990	AP007990 Lotus cor	c 192	15	75.0	194834	9	AC119864	AC119864 Mus muscu
c 120	15	75.0	62198	5	AJ621739	AJ621739 Gallus ga	c 193	15	75.0	196208	9	AC134426	AC134426 Mus muscu
c 121	15	75.0	63955	14	AC135981	AC135981 Homo sapi	c 194	15	75.0	197843	14	AC093139	AC093139 Pan trogl
c 122	15	75.0	68620	2	AC004301	AC004301 Drosophil	c 195	15	75.0	197870	14	AC091075	AC091075 Homo sapi
c 123	15	75.0	82261	8	AC000065	AC000065 Homo sapi	c 196	15	75.0	199208	14	AC019236	AC019236 Homo sapi
c 124	15	75.0	99899	14	AC107170	Continuation (4 of	c 197	15	75.0	204091	14	AC152442	AC152442 Bos tauru
c 125	15	75.0	110000	1	BA000035	Continuation (13 o	c 198	15	75.0	206665	14	AC092872	AC092872 Pan trogl
c 126	15	75.0	110000	1	CP000057	Continuation (12 o	c 199	15	75.0	208353	14	AC128744	AC128744 Rattus no
c 127	15	75.0	110000	6	BD426631	Continuation (12 o	c 200	15	75.0	208492	8	AC062033	AC062033 Homo sapi
c 128	15	75.0	110000	6	AR274513	Continuation (12 o	c 201	15	75.0	209421	8	AC090543	AC090543 Homo sapi
c 129	15	75.0	110000	6	AR632719	Continuation (12 o	c 202	15	75.0	210851	9	AC117686	AC117686 Mus muscu
c 130	15	75.0	110000	15	AE017347	Continuation (12 o	c 203	15	75.0	211288	5	AC159985	AC159985 Gallus ga
c 131	15	75.0	110000	15	AP008211	Continuation (195	c 204	15	75.0	212850	14	AC156553	AC156553 Mus muscu
c 132	15	75.0	110000	15	AP008213	Continuation (9 of	c 205	15	75.0	213476	14	AC114589	AC114589 Mus muscu
c 133	15	75.0	110060	14	RN339J16	Continuation (242	c 206	15	75.0	216470	14	AC119591	AC119591 Rattus no
c 134	15	75.0	110526	14	AC019598	CR937050 Rattus no	c 207	15	75.0	218133	14	AC130751	AC130751 Rattus no
c 135	15	75.0	110526	15	AP005486	AP005486 Oryza sat	c 208	15	75.0	218508	9	AC113497	AC113497 Mus muscu
c 136	15	75.0	113982	15	AP005486	CT009521 Medicago	c 209	15	75.0	218747	9	AC093339	AC093339 Mus muscu
c 137	15	75.0	116975	14	CT009521	AP003927 Oryza sat	c 210	15	75.0	219635	14	AC123204	AC123204 Rattus no
c 138	15	75.0	117240	14	AP003927	AC146024 Pan trogl	c 211	15	75.0	220803	8	AY128534	AY128534 Homo sapi
c 139	15	75.0	13263	14	AC126235	AL607079 Human DNA	c 212	15	75.0	220886	14	AC131001	AC131001 Rattus no
c 140	15	75.0	136107	15	AP003846	AC126235 Canis fam	c 213	15	75.0	221265	14	CT025587	CT025587 Mus muscu
c 141	15	75.0	136188	14	AC160111	AP003846 Oryza sat	c 214	15	75.0	221389	14	AC112122	AC112122 Rattus no
c 142	15	75.0	139452	9	RN476F3	CR937053 Rattus no	c 215	15	75.0	222207	14	AC128947	AC128947 Rattus no
c 143	15	75.0	139711	14	AC153093	AC153093 Ornithorh	c 216	15	75.0	222759	14	AC073177	AC073177 Homo sapi
c 144	15	75.0	139947	8	AC146024	AL607079 Human DNA	c 217	15	75.0	224866	14	AC107542	AC107542 Rattus no
c 145	15	75.0	144569	8	AL607079	AC105755 Homo sapi	c 218	15	75.0	226078	14	AC124018	AC124018 Rattus no
c 146	15	75.0	145673	8	AC105755	CR545470 Danio rer	c 219	15	75.0	226112	14	AC162526	AC162526 Mus muscu
c 147	15	75.0	145871	14	CR545470	AC161289 Pan trogl	c 220	15	75.0	228375	14	AC114460	AC114460 Rattus no
c 148	15	75.0	151213	8	AC161289	AL009172 Human DNA	c 221	15	75.0	229212	14	AC114168	AC114168 Rattus no
c 149	15	75.0	156601	8	HS212G6	AC120991 Oryza sat	c 222	15	75.0	229494	14	AC097740	AC097740 Rattus no
c 150	15	75.0	157069	15	AC120991	AC114679 Mus muscu	c 223	15	75.0	232077	14	AC091348	AC091348 Rattus no
c 151	15	75.0	157152	9	AC114679	AC008244 Homo sapi	c 224	15	75.0	232451	14	AC103060	AC103060 Rattus no
c 152	15	75.0	157216	14	AC008244	AC145492 Rattus no	c 225	15	75.0	233721	14	AC010159	AC010159 Homo sapi
c 153	15	75.0	157789	14	AC145492	AC149980 Strongylo	c 226	15	75.0	234117	14	AC130985	AC130985 Rattus no
c 154	15	75.0	159489	14	AC149980	AC006480 Homo sapi	c 227	15	75.0	235075	14	AC137044	AC137044 Rattus no
c 155	15	75.0	159752	8	AC006480	AC133327 Rattus no	c 228	15	75.0	235897	14	AC095685	AC095685 Rattus no
c 156	15	75.0	162167	14	AC133327	AC18771 Homo sapi	c 229	15	75.0	236039	14	AC121212	AC121212 Rattus no
c 157	15	75.0	162572	14	AC018377	AC160502 Bos tauru	c 230	15	75.0	236112	14	AC132015	AC132015 Rattus no
c 158	15	75.0	162581	14	AC121658	AL953905 Zebrafish	c 231	15	75.0	236333	9	AC138277	AC138277 Bos tauru
c 159	15	75.0	163505	14	AC160502	AC162880 Mus muscu	c 232	15	75.0	237286	14	AC095523	AC095523 Rattus no
c 160	15	75.0	163506	5	AL953905	AC117783 Mus muscu	c 233	15	75.0	237289	14	AC110965	AC110965 Rattus no
c 161	15	75.0	164201	9	AC162880	AC136809 Rattus no	c 234	15	75.0	238749	14	AC126156	AC126156 Rattus no
c 162	15	75.0	164296	9	AC117783	AC021899 Homo sapi	c 235	15	75.0	245139	14	AC103553	AC103553 Rattus no
c 163	15	75.0	164377	14	AC021899	AC106940 Rattus no	c 236	15	75.0	245659	14	AC106940	AC106940 Rattus no
c 164	15	75.0	164468	14	AC021899		c 237	15	75.0	245845	14	AC106940	AC106940 Rattus no





384	14	70.0	3140	6	CS041632 Sequence	457	14	70.0	42017	14	AC147080
385	14	70.0	3160	8	BC045641 Homo sapi	c 458	14	70.0	45349	8	AC011525
386	14	70.0	3295	9	MUSF1BLP2	c 459	14	70.0	45626	14	AY207045
387	14	70.0	3349	15	AK066418	c 460	14	70.0	48787	15	DQ115393
388	14	70.0	3410	15	SCYNL085W	c 461	14	70.0	50701	1	CP000082_26
389	14	70.0	3702	9	RNU66878	c 462	14	70.0	50927	6	CQ869661
390	14	70.0	3803	15	AK065064	c 463	14	70.0	52284	8	CH19R25187
391	14	70.0	4152	6	AR450996	c 464	14	70.0	53157	8	AL136302
392	14	70.0	4169	6	C0612380	c 465	14	70.0	54425	8	AC095352
393	14	70.0	4278	5	BC058830	c 466	14	70.0	58348	14	AC144883
394	14	70.0	4350	15	AF432498	c 467	14	70.0	60886	8	AL591969
395	14	70.0	4427	8	AB208916	c 468	14	70.0	61021	14	AC101916
396	14	70.0	4539	6	AR576998	c 469	14	70.0	62370	14	AL451008
397	14	70.0	4539	6	AX135962	c 470	14	70.0	62377	14	AL360009
398	14	70.0	4539	6	BD015022	c 471	14	70.0	63062	14	AL360009
399	14	70.0	4617	8	HSU80918	c 472	14	70.0	63062	14	AC135057
400	14	70.0	4678	8	HSU80917	c 473	14	70.0	63981	14	AC139146
401	14	70.0	5030	6	CQ591719	c 474	14	70.0	63981	14	AC139146
402	14	70.0	5358	6	CQ585377	c 475	14	70.0	65194	14	AC111098
403	14	70.0	5403	6	BD062061	c 476	14	70.0	66082	14	AC109316
404	14	70.0	5403	6	AR212643	c 477	14	70.0	66129	8	AP000281
405	14	70.0	5403	6	AR590684	c 478	14	70.0	67046	8	HSU63721
406	14	70.0	5403	9	AF025818	c 479	14	70.0	67598	14	AC124282
407	14	70.0	5419	14	AC018287	c 480	14	70.0	68491	5	AJ621744
408	14	70.0	5502	6	AR198316	c 481	14	70.0	68548	8	HSU70984
409	14	70.0	5502	6	AR648968	c 482	14	70.0	68984	14	AC101203
410	14	70.0	5865	8	HS1ZFI	c 483	14	70.0	69411	15	AP004898
411	14	70.0	6827	13	AF238884	c 484	14	70.0	74007	14	AC101438
412	14	70.0	7356	15	AY1919830	c 485	14	70.0	75430	14	AC166182
413	14	70.0	8002	6	CQ585590	c 486	14	70.0	76177	14	AC021620
414	14	70.0	8438	4	BTY17260	c 487	14	70.0	77138	14	AC025299
415	14	70.0	10271	1	U32790	c 488	14	70.0	80000	6	CQ819404
416	14	70.0	10271	6	AX191761	c 489	14	70.0	82288	14	AC164698
417	14	70.0	10554	1	AE006778	c 490	14	70.0	82360	2	AC084455
418	14	70.0	10819	1	AE007352	c 491	14	70.0	82510	15	AC146328
419	14	70.0	10834	1	AE004020	c 492	14	70.0	82818	14	AC012754
420	14	70.0	10949	1	AE008416	c 493	14	70.0	83684	8	AC105290
421	14	70.0	10966	1	AE015527	c 494	14	70.0	84415	14	AP008177
422	14	70.0	11093	1	AE010132	c 495	14	70.0	84972	8	AL365499
423	14	70.0	11309	1	AE004051	c 496	14	70.0	85694	15	BSX42599
424	14	70.0	11647	14	AC153648_5	c 497	14	70.0	88823	8	AL713965
425	14	70.0	14095	8	AC135052	c 498	14	70.0	89328	6	AX329823
426	14	70.0	14283	6	CQ587180	c 499	14	70.0	89328	8	HSJ398C2
427	14	70.0	15564	1	AB185495	c 500	14	70.0	94835	8	HSJ398C2
428	14	70.0	16995	6	CQ788991	c 501	14	70.0	96013	8	AL353779
429	14	70.0	16995	6	AR218850	c 502	14	70.0	96597	6	AX695611
430	14	70.0	16995	6	BD003762	c 503	14	70.0	97641	14	AP007792
431	14	70.0	17933	15	SCORFSDNA	c 504	14	70.0	98188	8	AC005056
432	14	70.0	20866	8	HSU62292	c 505	14	70.0	98935	8	AC004935
433	14	70.0	23743	1	AP005225	c 506	14	70.0	99824	14	AC157722
434	14	70.0	24526	6	AX472617	c 507	14	70.0	100000	8	AP000040
435	14	70.0	24965	14	AC014507	c 508	14	70.0	100000	8	AP000108
436	14	70.0	26130	1	AF212041	c 509	14	70.0	100000	8	AP000184
437	14	70.0	27856	1	CG1420072	c 510	14	70.0	100268	14	AC151361
438	14	70.0	31986	15	AF458976	c 511	14	70.0	101172	15	AP006364
439	14	70.0	32051	15	AF458978	c 512	14	70.0	101442	8	AC151340
440	14	70.0	32053	15	AF458981	c 513	14	70.0	104135	8	AC113373
441	14	70.0	32054	15	AF458969	c 514	14	70.0	104729	8	HS118B18
442	14	70.0	32055	15	AF458970	c 515	14	70.0	107120	14	AC091831
443	14	70.0	32055	15	AF458975	c 516	14	70.0	107482	15	BSX08789
444	14	70.0	32055	15	AF458977	c 517	14	70.0	108836	8	AC069137
445	14	70.0	32056	15	AF458980	c 518	14	70.0	110000	1	CR522870_28
446	14	70.0	32057	15	AF458979	c 519	14	70.0	110000	1	CR522870_28
447	14	70.0	32077	15	AF458972	c 520	14	70.0	110000	1	CR626927_02
448	14	70.0	32077	15	AF458974	c 521	14	70.0	110000	1	CR626927_25
449	14	70.0	32078	15	AF458971	c 522	14	70.0	110000	1	AB008692_02
450	14	70.0	32078	15	AF458973	c 523	14	70.0	110000	1	AB009951_08
451	14	70.0	34971	8	AP006749	c 524	14	70.0	110000	1	AE016853_37
452	14	70.0	35202	15	AC149413	c 525	14	70.0	110000	1	AE017340_20
453	14	70.0	35484	2	CET26E4	c 526	14	70.0	110000	1	AP006840_01
454	14	70.0	39539	14	AC145113	c 527	14	70.0	110000	1	AP006841_02
455	14	70.0	40875	8	AP000702	c 528	14	70.0	110000	1	AP006841_24
456	14	70.0	41541	8	AP000703	c 529	14	70.0	110000	1	AP008231_16







968 14 70.0 254973 14 AC127437  
c 969 14 70.0 255326 14 AC136556  
970 14 70.0 255342 9 AC109197  
971 14 70.0 255463 2 AE003735  
972 14 70.0 255596 14 AC095918  
c 973 14 70.0 255632 14 AC078997  
974 14 70.0 255881 14 AC102994  
c 975 14 70.0 256266 14 AC096361  
c 976 14 70.0 256584 14 AC133754  
977 14 70.0 257767 14 AC095974  
c 978 14 70.0 258328 14 AC133789  
979 14 70.0 258769 14 AC112745  
c 980 14 70.0 258939 14 AC110694  
c 981 14 70.0 259900 14 AC097550  
982 14 70.0 260421 14 AC152288  
983 14 70.0 260673 2 AE003612  
984 14 70.0 262061 14 AC117172  
985 14 70.0 262090 14 AC106054  
c 986 14 70.0 263405 14 AC124845  
c 987 14 70.0 264975 14 AC134732  
988 14 70.0 265537 14 AC087228  
c 989 14 70.0 265915 14 AC111228  
990 14 70.0 266851 2 AE003813  
c 991 14 70.0 267392 14 AC099071  
992 14 70.0 267512 14 AC095228  
c 993 14 70.0 267795 14 AC096390  
c 994 14 70.0 268894 14 AC129448  
c 995 14 70.0 270222 14 AC093976  
c 996 14 70.0 271043 14 AC164033  
c 997 14 70.0 271080 14 AC123123  
998 14 70.0 274560 14 AC099390  
c 999 14 70.0 275566 14 AC112796  
c1000 14 70.0 278514 14 AC095674

## ALIGNMENTS

\* RESULT 1  
LOCUS AR576665/c 254 bp DNA linear PAT 14-DEC-2004  
DEFINITION Sequence 12 from patent US 6777182.  
ACCESSION AR576665  
VERSION AR576665.1 GI:56578950  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 254)  
AUTHORS Baban,S., Bernard,M., Cherry,E., Gosselin,D., Hugo,P., Malette,B.,  
Miron,P., Prive,C. and Shazand,K.  
TITLE Methods for determining the likelihood of endometriosis in a female  
JOURNAL Subject  
Patent: US 6777182-A 12 17-AUG-2004;  
Metriogene Biosciences Inc.; Montreal;  
CAX;  
FEATURES  
source Location/Qualifiers  
1..254  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 254;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20  
|||||  
Db 90 GCTTTGGTTGGGCAACACAT 71

RESULT 2  
AX229886/c

LOCUS AX229886 254 bp RNA linear PAT 11-SEP-2001  
DEFINITION Sequence 12 from Patent WO0162959.  
ACCESSION AX229886  
VERSION AX229886.1 GI:15592021  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Baban,S., Bernard,M., Cherry,E., Gosselin,D., Hugo,P., Malette,B.,  
Miron,P., Prive,C. and Shazand,K.  
TITLE Endometriosis-related markers and uses thereof  
JOURNAL Patent: WO 0162959-A 12 30-AUG-2001;  
Procera Biosciences Inc. (CA)  
FEATURES  
source Location/Qualifiers  
1..254  
/organism="Homo sapiens"  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 254;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20  
|||||  
Db 90 GCTTTGGTTGGGCAACACAT 71

RESULT 3  
CQ679217/c 303 bp DNA linear PAT 03-FEB-2004  
LOCUS CQ679217  
DEFINITION Sequence 24143 from Patent WO02070737.  
ACCESSION CQ679217  
VERSION CQ679217.1 GI:42191349  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1  
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.  
TITLE Compositions and methods relating to osteoarthritis  
JOURNAL Patent: WO 02070737-A 24143 12-SEP-2002;  
Chondrogene Inc. (CA)  
FEATURES  
source Location/Qualifiers  
1..303  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 303;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20  
|||||  
Db 184 GCTTTGGTTGGGCAACACAT 165

RESULT 4  
CQ690723/c 447 bp DNA linear PAT 03-FEB-2004  
LOCUS CQ690723  
DEFINITION Sequence 35649 from Patent WO02070737.  
ACCESSION CQ690723  
VERSION CQ690723.1 GI:42228057  
KEYWORDS Homo sapiens (human)  
SOURCE

ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS Liaw,C.C., Marshall,W.E. and Zhang,H.  
TITLE Compositions and methods relating to osteoarthritis  
JOURNAL Patent: WO 02070737-A 35649 12-SEP-2002;  
Chondrogene Inc. (CA)

FEATURES  
source Location/Qualifiers  
1. .447  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 447;  
Best Local Similarity 100.0%; Pred. No. 0.36; Indels 0; Gaps 0;  
Matches 20; Conservative 0; Mismatches 0

Qy 1 GCTTTGGTGGCAACACAT 20  
Db 95 GCTTTGGTGGCAACACAT 76

RESULT 5  
AX778044/c  
LOCUS AX778044 578 bp DNA linear PAT 14-JUL-2003  
DEFINITION Sequence 201 from Patent WO03039443.  
ACCESSION AX778044  
VERSION AX778044.1 GI:32695038  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,  
Dugas,M., Ellis,R., Brors,B. and Mergenthaler,S.  
TITLE Novel genetic markers for leukemias  
JOURNAL Patent: WO 03039443-A 201 15-MAY-2003;  
Deutsches Krebsforschungszentrum (DE);  
Ludwig-Maximilian-Universitaet Muenchen (DE); Haferlach, Torsten,  
PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)

FEATURES  
source Location/Qualifiers  
1. .578  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 578;  
Best Local Similarity 100.0%; Pred. No. 0.35; Indels 0; Gaps 0;  
Matches 20; Conservative 0; Mismatches 0

Qy 1 GCTTTGGTGGCAACACAT 20  
Db 437 GCTTTGGTGGCAACACAT 418

RESULT 6  
BT007455/c  
LOCUS BT007455 1968 bp mRNA linear PRI 13-MAY-2003  
DEFINITION Homo sapiens forkhead box O1A (rhabdomyosarcoma) mRNA, complete  
cds.  
ACCESSION BT007455  
VERSION BT007455.1 GI:30583748  
KEYWORDS FLI CDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.  
1 (bases 1 to 1968)  
Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,  
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,  
Phelan,M. and Farmer,A.  
TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor  
vector  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1968)  
AUTHORS Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,  
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,  
Phelan,M. and Farmer,A.  
TITLE Direct Submision  
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow  
Circle, Palo Alto, CA 94303, USA  
COMMENT This CDS clone is a part of a collection of human full length  
expression clones generated by BD Biosciences Clontech and the  
Harvard Institute of Proteomics. Each CDS has been cloned in two  
forms: with and without stop-codon (to allow fusion with C-terminal  
tag). The CDS has been directionally cloned using BD In-Fusion(TM)  
cloning system between the SalI and HindIII sites of the pDNR-DUAL  
vector. Additional sequences in the clone: 'ACC' after SalI site  
and before 'ATG' to provide Kozak consensus sequence; 'GG' after  
last codon and before HindIII site to maintain reading frame.  
Clone distribution: http://bioinfo.clontech.com/orfclones.

FEATURES  
source Location/Qualifiers  
1. .1968  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GH00949X1.0"  
/clone\_lib="BD Creator(TM) CDS Library derived from MGC  
collection"  
/lab\_host="DH5alpha T1 resistant"  
/note="Vector: pDNR-Dual"  
1. .1968  
/codon\_start=1  
/product="forkhead box O1A (rhabdomyosarcoma)"  
/protein\_id="AAP36123.1"  
/db\_xref="GI:30583749"  
/translation="MARAPQVVEIDPDPELPRPESCTWPLPRPFSOSNATSPPAP  
SGSAAMPDAAGLPSAASAAVADPMNLSESEDFPQAPGSAVAATAAAAAA  
TGGCGDFQGPENGALHPAPPQPPPPGLSPQHPVPPAAAGPLAGQPRKSSRRNAW  
GNLSADLITKAIESSAEKRLTLSQIYEMWYKVPYFKGSDNSAGMKNIRHNL  
LHSEFIRVQNETGKSWMLNPGGSGKSPRRRAAAMNNKSNASTISGRSLPTKAS  
LQSGGEGADSPGQSKWPASPGSHNDPFDNWSTFRPTSSNASTISGRSLPTKAS  
ODDLGEGDVHSMVYPPSAAKWASTLPSLSEISNENMENLADNLNLSPTSLTSTQ  
SSPCTMMQQTFCYFAPPTSLNSPSPNYQKYQSSMSPLPQNPQTLQDNKSSYG  
GMSQYNCAPLGLLKSLTSDSPPHNDIMTPDPGVAQPNRVLGQNMVMPNSVMTYV  
SQASHNKNMNPSSHTPHGAQQTSAVNGRPLPTVSTVMTPTSGMNRLLTQVKTVPQVPL  
PHPMQMSALGGYSSVSCNGYGRMGLLHQEKLPDLQGMFIERLDCDMESIIRNDLMD  
GDTLDFNFDNVLNQSPPHVSVKTTTTHSVSG"

ORIGIN  
Query Match 100.0%; Score 20; DB 8; Length 1968;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTGGCAACACAT 20  
Db 1920 GCTTTGGTGGCAACACAT 1901

RESULT 7  
AY890173/c  
LOCUS AY890173 1968 bp mRNA linear SYN 29-MAR-2005  
DEFINITION Synthetic construct Homo sapiens clone FLH015808.01X forkhead box  
O1A (FOXO1A) mRNA, complete cds.  
ACCESSION AY890173  
VERSION AY890173.1 GI:61361857  
KEYWORDS Human ORF Project.  
SOURCE synthetic construct  
ORGANISM synthetic construct

other sequences; artificial sequences. 1. (bases 1 to 1968) Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F., Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E., Williamson,J. and LaBaer,J. Cloning of human full-length CDS in Creator (TM) recombinational vector system Unpublished 2. (bases 1 to 1968) Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F., Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E., Williamson,J. and LaBaer,J. Direct Submission Submitted (04-JAN-2005) Biological Chemistry and Molecular Pharmacology, Harvard Institute of Proteomics, 320 Charles St., Cambridge, MA 02141, USA This CDS clone is a part of a collection of human full-length expression clones generated by Harvard Institute of Proteomics. This ORF clone has been cloned with normalized stop-codon. The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the Sali and HindIII sites of the pDNR-Dual vector. Additional sequences in the clone: 'ACC' after Sali site and before 'ATG' to provide Kozak consensus sequence. Each clone is clonally isolated and full-length sequence-verified.	synthetic construct synthetic construct other sequences; artificial sequences. 1. (bases 1 to 1968) Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F., Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E., Williamson,J. and LaBaer,J. Cloning of human full-length CDS in Creator (TM) recombinational vector system Unpublished 2. (bases 1 to 1968) Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F., Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E., Williamson,J. and LaBaer,J. Direct Submission Submitted (04-JAN-2005) Biological Chemistry and Molecular Pharmacology, Harvard Institute of Proteomics, 320 Charles St., Cambridge, MA 02141, USA This CDS clone is a part of a collection of human full-length expression clones generated by Harvard Institute of Proteomics. This ORF clone has been cloned with normalized stop-codon. The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the Sali and HindIII sites of the pDNR-Dual vector. Additional sequences in the clone: 'ACC' after Sali site and before 'ATG' to provide Kozak consensus sequence. Each clone is clonally isolated and full-length sequence-verified.
REFERENCE AUTHORS	ORGANISM
TITLE	REFERENCE
JOURNAL REFERENCE AUTHORS	AUTHORS
TITLE	TITLE
JOURNAL	JOURNAL
COMMENT	COMMENT
FEATURES	FEATURES
source	source
gene	gene
CDS	CDS
ORIGIN	ORIGIN
Query Match Best Local Similarity Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Query Match Best Local Similarity Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCTTTGGTTGGCAACACAT 20 	Qy 1 GCTTTGGTTGGCAACACAT 20 
Db 1920 GCTTTGGTTGGCAACACAT 1901	Db 1920 GCTTTGGTTGGCAACACAT 1901
RESULT 8 AY890174/c LOCUS DEFINITION ACCSSION VERSION KEYWORDS	RESULT 9 AF126056/c LOCUS DEFINITION ACCSSION VERSION



KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Nakae, J., Park, B.C. and Accili, D.  
TITLE Insulin stimulates phosphorylation of the forkhead transcription factor FOXO1 on serine 253 through a Wortmannin-sensitive pathway  
JOURNAL J. Biol. Chem. 274 (23), 15982-15985 (1999)  
PUBMED 10347145  
REFERENCE  
AUTHORS Nakae, J. and Accili, D.  
TITLE Direct Submission  
JOURNAL Submitted (04-FEB-1999) Developmental Endocrinology Branch, National Institutes of Health, Building 10, Room 10D18, Bethesda, MD 20892, USA

FEATURES  
source  
1..2103  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
gene  
1..2103  
/gene="Ptkhr"  
CDS  
7..1965  
/gene="Ptkhr"  
/function="transcription factor"  
/codon\_start=1  
/product="forkhead protein FOXH"  
/protein\_id="AAD40636.1"  
/db\_xref="GI:5163279"  
/translation="MAEAPQVVEIDPDEPLPRQSCWTWPLPRPFNSQNSATSSPAP SGGAANPDAAAGLPSASAAVDTFMSNLSEESDFARFCVCVAAAAAASRLGCG DFQPGACGCHPAPQPPPTGSLQSPVPVPSAAAAGPLAGQPKTSSRRNAGNL SYADLITKAISSAEKRLTLOIYEMVKSVFYFKDGDNSAGKWSIRHNSLHLS KFIQVNGSTGCKSWMLNPEGSKGSPRRRAASMDNNKSKFAKSGRAAKGASLQS QGEGDPSGQFSQKWPASPGSHNDPDFNWSFTFRPTSSNASTISGRSLPMTQDD LGDGVSLVPPPSAAKWAATLPSLSEISNPNMENLDNLNLSSTSLTSTVQSSSL GMMQQTFCYFAPATSLNPSFNPNYKYTYGQSSMSPLPQPMQTLQDSKSYGGLN QYNCAAGLLKELLTSDSPHNDIMSPDVPDGAQNSRVLGQVMMGNSVMPGSGQA SHKMWNPSSHTPGHQAQTASVNGRTLPHVVTWPTHTSAMNRLTPVKTPLQVPLSHP MNKMSALGSYSSVSCNGYGRMGVLHOKLPSDLDMGFIERLDCDMESIIRNDPMDGDT LDFNPDVLPNQSPFHSVKTTHSWVSG"

ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 2103;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 1917 GCTTTGGTTGGCAACACAT 1898  
|||||

RESULT 10  
AY094061/c  
LOCUS  
DEFINITION Sus scrofa forkhead/winged helix transcription factor FOXO1a  
ACCESSION AY094061  
VERSION AY094061.1 GI:20384962  
KEYWORDS  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa

REFERENCE  
AUTHORS Zhu, Q., Cunningham, M.A. and Hammond, J.M.  
TITLE FKHR Expression in Porcine Granulosa Cells  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2162)

AUTHORS Zhu, Q., Cunningham, M.A. and Hammond, J.M.  
TITLE Direct Submission  
JOURNAL Submitted (02-APR-2002) Dept. of Medicine/Div. of Endocrinology, Pennsylvania State Univ. College of Medicine, 500 University Dr., Hershey, PA 17033, USA

FEATURES  
source  
1..2162  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
gene  
1..2162  
/gene="foxola"  
CDS  
33..2021  
/gene="foxola"  
/codon\_start=1  
/product="forkhead/winged helix transcription factor FOXO1a"  
/protein\_id="AAM19156.1"  
/db\_xref="GI:20384963"  
/translation="MAEAPQVVEIDPDEPLPRQSCWTWPLPRPFNSQNSATSSPAP SGGAANPDAAAGLPSASAAVDTFMSNLSEESDFARFCVCVAAAAAASRLGCG DFQPGACGCHPAPQPPPTGSLQSPVPVPSAAAAGPLAGQPKTSSRRNAGNL SYADLITKAISSAEKRLTLOIYEMVKSVFYFKDGDNSAGKWSIRHNSLHLS KFIQVNGSTGCKSWMLNPEGSKGSPRRRAASMDNNKSKFAKSGRAAKGASLQS QGEGDPSGQFSQKWPASPGSHNDPDFNWSFTFRPTSSNASTISGRSLPMTQDD LGDGVSLVPPPSAAKWAATLPSLSEISNPNMENLDNLNLSSTSLTSTVQSSSL GMMQQTFCYFAPATSLNPSFNPNYKYTYGQSSMSPLPQPMQTLQDSKSYGGLN QYNCAAGLLKELLTSDSPHNDIMSPDVPDGAQNSRVLGQVMMGNSVMPGSGQA SHKMWNPSSHTPGHQAQTASVNGRTLPHVVTWPTHTSAMNRLTPVKTPLQVPLSHP MNKMSALGSYSSVSCNGYGRMGVLHOKLPSDLDMGFIERLDCDMESIIRNDPMDGDT LDFNPDVLPNQSPFHSVKTTHSWVSG"

ORIGIN  
Query Match 100.0%; Score 20; DB 4; Length 2162;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 1973 GCTTTGGTTGGCAACACAT 1954  
|||||

RESULT 11  
BC021981/c  
LOCUS  
DEFINITION Homo sapiens forkhead box O1a (rhabdomyosarcoma), mRNA (cDNA clone MGC:1750 IMAGE:2959021), complete cds.  
ACCESSION BC021981  
VERSION BC021981.2 GI:33869892  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2413)

REFERENCE  
AUTHORS Klausner, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G., Stauber, R.D., Collins, F.S., Wagner, L., Sherman, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huly, K.S.W., Villalón, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalska, U., Smalley, D.E.,

CONSRMT  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
CONSRMT  
TITLE  
JOURNAL  
REMARK  
COMMENT

Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Mammalian Gene Collection Program Team  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 2413)

NIH MGC Project  
Direct Submission  
Submitted (22-JAN-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Aug 19, 2003 this sequence version replaced gi:18314374.  
Contact: MGC help desk  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadan@systemsbio.org](mailto:amadan@systemsbio.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 2 Row: 0 Column: 10.

FEATURES  
source  
Location/Qualifiers  
1. .2413  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:1750 IMAGE:2959021"  
/tissue\_type="Lymph, Burkitt lymphoma"  
/clone\_id="NIH\_MGC\_8"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
1. .2413  
/gene="FOXO1A"  
/note="synonyms: FOXO1, FKH1"  
/db\_xref="GeneID:2308"  
314. .2281  
/db\_xref="MIM:136533"  
/gene="FOXO1A"  
/codon\_start=1  
/product="forkhead box O1a"  
/protein\_id="AAH21981.1"  
/db\_xref="GI:18314375"  
/db\_xref="GeneID:2308"  
/db\_xref="MIM:136533"

RESULT 12  
AF114258/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
gene  
5'UTR  
CDS  
3'UTR  
ORIGIN  
Query Match  
Best Local Similarity  
Matches  
Qy  
Db  
RESULT 13  
BC070065/c  
LOCUS  
DEFINITION

AF114258 2460 bp mRNA linear ROD 06-JUN-2001  
Mus musculus forkhead protein FKHR1 (Fkhr1) mRNA, complete cds.  
AF114258  
AF114258.1 GI:5348329  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
1 (bases 1 to 2460)  
Biggs,W.H. III and Cavenee,W.K.  
Identification and characterization of members of the FKHR (FOX O)  
subclass of winged-helix transcription factors in the mouse  
Mamm. Genome 12 (6), 416-425 (2001)  
11353388  
2 (bases 1 to 2460)  
Biggs,W.H. III, Cavenee,W.K. and Arden,K.C.  
Direct Submission  
Submitted (15-DEC-1998) Ludwig Institute for Cancer Research, San  
Diego Branch, 9500 Gilman Drive, La Jolla, CA 92093-0660, USA  
Location/Qualifiers  
1. .2460  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/chromosome="3"  
/dev\_stage="E8.5 embryo"  
1. .>2460  
/gene="Fkhr1"  
1. .383  
/gene="Fkhr1"  
384. .2342  
/gene="Fkhr1"  
/note="member of FKHR-subclass forkhead/winged helix  
family"  
/codon\_start=1  
/product="forkhead protein FKHR1"  
/protein\_id="AAB42106.1"  
/db\_xref="GI:5348330"  
/translation="MAEAPQVETDPEPLPRQSCWTWPLPRPFQNSNTSSPAP  
SGSAAANPDAAAGLPSAASAVSDTFMSNLILLESDPAPAGCVAAVAAAAAGRLCG  
DFQGPAGCVHPAPQPPTGSLQHPVPPAAAGPLAGQPRKTSSSRRNAGNL  
SYADLTKALESSEAKRLTSLQIYEMVKSVPFKDKGDSNSAGWKNIRHNLSLHS  
KFIRVQNEGTSKSSWWMLNPEGSGSGSPRRRAASMDNNKFAKSRGAARAKKASLOS  
GQEGDVSLSVYPPSAKMASTLPSLSSEISNPENMENLLNLILSSPTSLTSTVQSSP  
LGDDVHSLVYPPSAKMASTLPSLSSEISNPENMENLLNLILSSPTSLTSTVQSSP  
GSMWQTPCYSFAPNTSLNSPSPNYKYTYGSSMSPLQPMPTQIDKSKSYGGLN  
OYNCAPLLKELLTSDSPPHNDIMSVPDGVAPQNSRVLGQNVMMGPNVMPAYGSOA  
SHNNMNPSSHTPHGHAQQTASVNGRTTPPHVNMTPHTSANNRLTPVKYTPLOVPLSRP  
MQMSALGQYSVSSCGYGRMGLHQEKLPSDLQGMPIERLDCDMSIIRNDLMDGT  
LDNFNFDNVLNQSFPFHSVKTTTTHSWVG"  
2345. .>2460  
/gene="Fkhr1"

100.0%; Score 20; DB 9; Length 2460;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCCTTGGTGGCAACACAT 20  
|||||  
2294 GCCTTGGTGGCAACACAT 2275

100.0%; Score 20; DB 8; Length 2413;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCCTTGGTGGCAACACAT 20  
|||||  
2233 GCCTTGGTGGCAACACAT 2214

```
ACCESSION BC070065
VERSION BC070065.1 GI:47123315
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 2482)
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shermen,C.M., Schuler,G.D.,
Altshuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusik,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Sheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettner,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2482)
NIH MGC Project
Direct Submission
Submitted (10-MAY-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAP Plate: 168 Row: n Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 9257221.
FEATURES
source
location/Qualifiers
1..2482
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30345006"
/tissue_type="Placenta, normal"
/clone_lib="NIH MGC_147"
/lab_host="DH10B"
/note="Vector: pBluescriptR"
<1..2482
/gene="FOXO1a"
/note="synonyms: FOXO1, FKH1"
/db_xref="GeneID:2308"
/db_xref="MIM:136533"
<1..2344
gene
CDS
/translation="MAEAPQVVEIDPPELPRLPRSCCTWPLPRPFSSNATSSPAP
SGCATPNFASAGLPAAAVASADFMNSLSLEESDFPQAPGSAVAAAATAAAT
GGLCDFQGLEAGCLHPAPPQPPGFLPSQHPVPPAAGPLAGQPRKSSRRNAWG
```

```
/gene="FOXO1a"
/codon_start=2
/product="FOXO1a protein"
/protein_id="AAH70065.2"
/db_xref="GI:71052130"
/db_xref="GeneID:2308"
/db_xref="MIM:136533"
/translation="HSTGSSAAGPLGRASGPRRPSVLPSSAALSAGARRRLCPGPAAL
AGRPVRAADPEPRCGWPREVKFWARASTPRLPSPFLPAAFPFOISDRFPAPP
RPPVLRSPPLGSPAAGGAGVTAEPVVEIDPPELPRLPRSCCTWPLPRPFSSQ
NSAASAPGSAAGANPDAAAGLPSAASAAVSADFMNSLSLEESDFPQAPGSAVAA
VAAAAAATGCGDFGPRAGCLHPAPPQPPGFLPSQHPVPPAAGPLAGQPRK
SSSRNAGNLSYADLLTYKALSSAEKRLTSLQIYEMVKSVYFPKDGDSNSGHW
KNSRNLSLSKFTIRVQNGTGGSSWMLAPEGGSGKSPRRRAASDNNNSKFAKSR
SRAKKASLGQGGEGDFSGQSPKWPASPGSHSDDFDNWFTRPTSTNSAATIS
GRLSPTMTQDGLGDMHMYPPPSAAKMASTLPSLSISINPENMLLDNLILSS
TQTSIDTQSGFTMMQTCVSPAPNTSLNSPSPNTQKYTYGSSMSPLPQMPLQT
LDKNKSSYGGMSQYNCAKGLKELLTSDSPHNDIMTPDVGVAQPNRSVLGNVMMG
PNSVMSTYGSQASHNKNWNPSSHTPHGAQQTSAVNGRLPLHTVTMTPTSGNRLITQ
VKTPQVPLPHPMQMSALGGYSSVSSCGYGRMGLLHQEKLPDLDGMPIERLDCDME
SIIRNDLMDGDTLDFNFDNLVLPNSFPHSVKTTTHSWVSG"
```

ORIGIN

Query Match 100.0%; Score 20; DB 8; Length 2482;  
Best Local Similarity 100.0%; Pred. No. 0.29; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0

QY 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 2296 GCTTTGGTTGGCAACACAT 2277  
|||||

RESULT 14  
AY255525/c  
LOCUS  
DEFINITION  
Spermophilus tridecemlineatus forkhead box O1a protein (FoxO1a)  
mRNA, complete cds.  
ACCESSION  
AY255525  
VERSION  
AY255525.1 GI:29373145  
KEYWORDS  
SOURCE  
ORGANISM  
Spermophilus tridecemlineatus (thirteen-lined ground squirrel)  
Spermophilus tridecemlineatus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Sciuridae; Xerinae; Marmotini; Spermophilus.  
REFERENCE  
1 (bases 1 to 2503)  
Cai, D., McCarron, R.M. and Hallenbeck, J.  
Cloning and characterization of a forkhead transcription factor  
gene, FoxO1a, from thirteen-lined ground squirrel  
Gene 343 (1), 203-209 (2004)  
2 (bases 1 to 2503)  
Cai, D. and Hallenbeck, J.M.  
Direct Submission  
Submitted (13-MAR-2003) NINDS, NTH, 9000 Rockville Pike, Bethesda,  
MD 20892, USA  
FEATURES  
source  
location/Qualifiers  
1..2503  
/organism="Spermophilus tridecemlineatus"  
/mol\_type="mRNA"  
/db\_xref="taxon:43179"  
/tissue\_type="brain"  
1..2503  
/gene="FoxO1a"  
398..2359  
/gene="FoxO1a"  
/note="transcription factor"  
/codon\_start=1  
/product="forkhead box O1a protein"  
/protein\_id="AAO72710.1"  
/db\_xref="GI:29373146"  
/translation="MAEAPQVVEIDPPELPRLPRSCCTWPLPRPFSSNATSSPAP  
SGCATPNFASAGLPAAAVASADFMNSLSLEESDFPQAPGSAVAAAATAAAT  
GGLCDFQGLEAGCLHPAPPQPPGFLPSQHPVPPAAGPLAGQPRKSSRRNAWG

LSYADLITKAIESSAEKRLTLISQIYEWVKSVPYFKDGDSSNSAGWKNIRHNLSLH  
SKFTRVQNEGTGKSSWMLNPEGGKSGKSPRRRAASMDNNSKFAKSRGQAOKKASIQ  
SGORGAGPSGQSKWPASPASHSDNDNDNWSFRPRTSSNASTISGRFSPIMTEOD  
DLGDDVHSLVYPPSASAKMASTLPSLSEINPENMENLNLNLLSPTSLTSTQSS  
PGSMQOQTPCYSFAPPSLSMSPFNQKTYTGSSSPSLSQMPQTLQDNKSSYGGI  
NQFNAQGLLKELLTSDSPPHNDIMSSVDPGVAQPNVGLGQNMVGPNLGNSSYPAYGQ  
ASHKNMNPSSHHPGHAQOTSVYNGRALPHVTNPMPTHTSGMRLTTPVKTPLOVPLIH  
HMOMSLGGYSSVSSCGYGRMGVLHQEKLPSDLDDGMFIERLDCDMESIIRNDLMDGD  
ALDPNFDNVLNQSPFPHGVKTTTHSWSG"

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 2503;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 2311 GCTTTGGTTGGCAACACAT 2292

RESULT 15  
AF178854/c

LOCUS AF178854 2725 bp mRNA linear SYN 27-DEC-1999  
DEFINITION Synthetic construct Pax3-forkhead fusion protein (Pax3/FKHR) mRNA,  
complete cds.

ACCESSION AF178854  
VERSION AF178854.1 GI:6636096

## KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 2725)

AUTHORS Lam, P. Y. P., Sublett, J. E., Hollenbach, A. D., Rousset, M. F.,

Epstein, J. A., Song, B., Lakkis, M. and Wang, C.

TITLE Author's Corrections

JOURNAL Mol. Cell. Biol. (2000) In press

REFERENCE 2 (bases 1 to 2725)

AUTHORS Naeye, C. W., Rakestraw, K., Hollenbach, A. D., Grosfeld, G. and

Rousset, M. F.

TITLE Direct Submission

JOURNAL Submitted (19-AUG-1999) Hartwell Center for Bioinformatics &

Biotechnology, St. Jude Children's Research Hospital, 332 N.

Lauderdale St., Memphis, TN 38105, USA

FEATURES Location/Qualifiers

1..2725

source /organism="synthetic construct"

/mol\_type="mRNA"

/db\_xref="taxon:32630"

/focus

1..630

source /organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/note="Pax3 sequence deposited in GenBank Accession Number

X59538"

631..2719

source /organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/note="derived from Pax3/FKHR fusion sequence deposited in

U02368"

gene 1..2725

/gene="Pax3/FKHR"

/note="compared to source sequence"

/replace="t"

209..2719

CDS /gene="Pax3/FKHR"

/note="transcription factor"

/codon\_start=1

/transl\_table=11

/product="Pax3-forkhead fusion protein"

/protein\_id="AAF20054.1"

/db\_xref="GI:6636097"  
/translation="MTTLAGAVPRMMRPGQNYPRSGFPLEVTPLQGRVNLGGV  
FNGRPLPNHRIKHIVEMAHGIRPCVISQLRVSHGCVSKILCRYQETGSIIRPAIG  
GSPKQVTTDPVEKKIIEYKREKPGMFWIRKLLKDAVCDRTNVTFSVSSILRLS  
KFGKEBEREADLERKEABESEKAKHSDIGILSERASAPQSDGSDIDSEPDLPKRR  
QRSRTTFTAEQLLEERAFERTHYPIDYITRELAQRAKLTEARVQVWFNRRARWK  
QAGANOLMAFNHLIPGFPPTAMPTLTPTQOLSETSYQPTSTPQASVPSLTVHRPQPL  
PPSTVHQSTTTPSNPDSSAYCLPSTRHGFSSYTDSPVPPSGPSNPMPTINGLSPON  
SIRNLSLHSEKFRVQNEGTGKSSWMLNPEGGKSGKSPRRRAASMDNNSKFAKSR  
AAKKSLSQGGAGDSGQSKWPASPASHSDNDNDNWSFRPRTSSNASTISGR  
LSPIMTEQDLDGEGDVHSMVYPPSAAKMASTLPSLSEINPENMENLNLNLSST  
SLVTSQSSPFTMQQPCYSFAPPTSLNSPSPNYQKTYTGSSMPLPQMPQTQLQ  
DNKSSYSGMSQYNCAKLLKELTSDSPPHNDIMTVPDGVAPNSVGLGQNMVGPN  
SVNSTVGSQASHKNMNPSSHTHPGHAQOTSAVNGRALPHVTNPMPTHTSGMRLTQVK  
TPVQVPLPHEPMQMSALGGYSSVSSCGYGRMGVLHQEKLPSDLDDGMFIERLDCDMESI  
IRNDLMDGDITLDFNFDNVLNQSPFPHGVKTTTHSWSG"

## ORIGIN

Query Match 100.0%; Score 20; DB 11; Length 2725;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 2671 GCTTTGGTTGGCAACACAT 2652

## RESULT 16

HSU02368/c

LOCUS HSU02368 3200 bp mRNA linear PRI 01-DEC-1993

DEFINITION Human PAX3/forkhead transcription factor gene fusion mRNA, complete

cds.

VERSION U02368

KEYWORDS U02368.1 GI:431253

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 3200)

AUTHORS Shapiro, D. N., Sublett, J. E., Li, B., Downing, J. R. and Naeye, C. W.

TITLE Fusion of PAX3 to a member of the forkhead family of transcription

factors in human alveolar rhabdomyosarcoma

Cancer Res. 53 (21), 5108-5112 (1993)

8221846

REFERENCE 2 (bases 1 to 3200)

AUTHORS Shapiro, D. N.

TITLE Direct Submission

JOURNAL Submitted (04-OCT-1993) Shapiro D. N., St. Jude Children's Research

Hospital, Experimental Oncology, 332 North Lauderdale, Memphis, TN

38105, USA

FEATURES Location/Qualifiers

1..3200

source /organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="2;13"

/map="2q35;q14"

/clone="pALV6.1"

/cell\_line="RH30"

/cell\_type="alveolar"

/tissue\_type="rhabdomyosarcoma"

/clone\_lib="lambda gt22A"

1..48

49..2559

5'UTR /codon\_start=1

CDS /product="PAX3 protein-forkhead transcription factor

fusion"

/protein\_id="AAC50053.1"

/db\_xref="GI:431254"

/translation="MTTLAGAVPRMMRPGQNYPRSGFPLEVTPLQGRVNLGGV

FNGRPLPNHRIKHIVEMAHGIRPCVISQLRVSHGCVSKILCRYQETGSIIRPAIG

GSKPKOVTTDPVEKKIEYKRENGMFSWIRDKLLKQAVCDNTVPVSSISRIILRS  
KFGKEEZEADLERKAESEKAKSHIDGILSERASAFQSDGSDISEDPLPKRK  
QKRSRTIFAEQEELEAFERKTHYPDIYIRELAQRAKLTEARQVQVFNFRKWRK  
QAGAQMFANHLIPGGFPPTAMPTLTFTYQLSYTSYQTSIPQAVSDPSTVHRPQPL  
PPSTHQSTIPNPDSSSAYCLPSTRHGFSTYDSFVPPSGSPNPNPTIGNGLSPON  
SIRHNLISLHSEKIRIVQNEGTGKSWMLNPEGKSGKSPRRRAASMDNNSKFAKSR  
AAKASLQSGGEGAGDSPGSQFSPKWPASPGSHNDNDPNDWSTFRPTSSNATISGR  
LSPITEQDLDGEGDVHSNVYPPSAKMASTLPSLSEISNPNMENLNDNLLSPPT  
SVTSTQSGFGTMMQOTPCYFAPNTSLNSPSPNYQKTYGCGSSMSPUPQMPITLQ  
DNKSYGSGMSQYNCAPGLLKELLTSDSPHNDIMTPVDGVAQPNRVLGQNMVGMGN  
SYMSTYGSQASHKMNWPSHSTPHQAQTSVNGRPLPHTVSTWPTSHGNNRLTQVK  
TPVQPLPHPMQMSALGYSSVSSCNGYGRMGLLHOKLPSLDLGDGMFIERLDCDMESI  
IRNDLMDGDTDFNFDNVLNQSPFHSVKTTTHSWVSG"  
148..544  
/note="paired box"  
/function="DNA binding domain"  
604..628  
/note="octapeptide"  
703..886  
/note="paired-type homeodomain"  
/function="DNA binding"  
1218..1371  
/note="forkhead domain"  
/function="DNA binding"  
2558..3200  
3'UTR  
ORIGIN  
Query Match 100.0%; Score 20; DB 8; Length 3200;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCTTTGGTTGGCAACACAT 20  
Db 2511 GCTTTGGTTGGCAACAT 2492  
RESULT 17  
CQ874011/c Q874011 3421 bp DNA linear PAT 27-SEP-2004  
LOCUS Sequence 430 from Patent WO2004076622.  
DEFINITION Q874011  
ACCESSION Q874011  
VERSION Q874011.1 GI:52747603  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Taira, K. and Kawasaki, H.  
TITLE Regulation of mammalian cells  
JOURNAL Patent: WO 2004076622-A 430 10-SEP-2004;  
National Institute of Advanced Industrial Science and Technology  
(JP)  
FEATURES  
source  
Location/Qualifiers  
1..3421  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 3421;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCTTTGGTTGGCAACACAT 20  
Db 1926 GCTTTGGTTGGCAACACAT 1907  
RESULT 18  
HSU02310/c HSU02310 3421 bp mRNA linear PRI 16-DEC-1993  
LOCUS

Human fork head domain protein (FKHR) mRNA, complete cds.  
U02310  
VERSION U02310.1 GI:435422  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3421)  
AUTHORS Galili, N.  
TITLE Fusion of a fork head domain to PAX-3 in the solid tumor alveolar  
rhabdomyosarcoma  
JOURNAL Nature Genetics 5(3), 230-235 (1993)  
REFERENCE 2 (bases 1 to 3421)  
AUTHORS Galili, N., Davis, R.J., Fredericks, W.J., Mukhopadhyay, S.,  
Rauscher, F.J. III, Emanuel, B.S., Rovera, G. and Barr, F.G.  
TITLE Direct Submission  
JOURNAL Submitted (01-OCT-1993) Naomi Galili, Wistar Institute, 3606 Spruce  
Street, Philadelphia, PA 19104 USA  
FEATURES  
source  
Location/Qualifiers  
1..3421  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..3421  
/gene="FKHR"  
7..1974  
/gene="FKHR"  
/codon\_start=1  
/product="fork head domain protein"  
/protein\_id="AAA03629.1"  
/translation="MAEAPQVVEIDPEFPRPSCWTPLPRPFSQSNSATSSPAP  
SGSAANPDAAAGLPSASAAVADFMNLSLLESEDFPOAPGSAVAATAAAAAA  
TGGLCGDPQGEAGCLHAPQPPPPGVSOHPVPPRAAGPLAGQPKSSSSRNAM  
GNLSYADLITKAIESSAEKRLTSLQIYEMWKSIFYKDKGNSNSAGWNSIRNLS  
LHSKPRVQNEGTGKSWMLNPEGKSGKSPRRRAASMDNNSKFAKSRRAAKKAS  
LQSGGEGAGDSPGSQFSPKWPASPGSHNDNDPNDWSTFRPTSSNATISGRISPMTE  
QDDLEGDVHSNVYPPSAKMASTLPSLSEISNPNMENLNDNLLSPPTIGNGLSPON  
SSPGTMMQOTPCYFAPNTSLNSPSPNYQKTYGCGSSMSPUPQMPITLQDNKSSYG  
GMSQYNCAPGLLKELLTSDSPHNDIMTPVDGVAQPNRVLGQNMVGMGNVMSVTVGL  
SQASHKMNWPSHSTPHQAQTSVNGRPLPHTVSTWPTSHGNNRLTQVKTPVQPL  
PHPMQMSALGYSSVSSCNGYGRMGLLHOKLPSLDLGDGMFIERLDCDMESIIRNDLMD  
GDTLDFNFDNVLNQSPFHSVKTTTHSWVSG"  
ORIGIN  
Query Match 100.0%; Score 20; DB 8; Length 3421;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCTTTGGTTGGCAACACAT 20  
Db 1926 GCTTTGGTTGGCAACAT 1907  
RESULT 19  
HSU02308/c HSU02308 3517 bp mRNA linear PRI 16-DEC-1993  
LOCUS Human PAX-3-FKHR gene fusion mRNA, partial cds.  
DEFINITION U02308  
ACCESSION U02308.1 GI:435418  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3517)  
AUTHORS Galili, N., Davis, R., Fredericks, W.J., Mukhopadhyay, S.,  
Rauscher, F.J. III, Emanuel, B.S., Rovera, G. and Barr, F.G.  
TITLE Fusion of a fork head domain gene to PAX3 in the solid tumor  
alveolar rhabdomyosarcoma

```
JOURNAL Nature Genetics 5(3), 230-235 (1993)
REFERENCE 2 (bases 1 to 3517)
AUTHORS Galili,N.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-1993) Naomi Galili, Wistar Institute, 3606 Spruce
Street, Philadelphia, PA 19104 USA
FEATURES
    source
        Location/Qualifiers
            1..3517
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                <1..2070
                    /note="PAX-3-FKHR gene fusion"
                    /codon_start=1
                    /protein_id="AAA03627.1"
                    /db_xref="GI:435419"
                    /translation="PSVSSISRLRSFKGEEBAAERKEASESEKKAHSDIGIL
SERASQDSGSDIDSEPDLPKARKQRKRTTFAEQLEEHVAFERTHYDPIYR
BELAQAKLTARVQVWFNSRRARWRQAGANQLMFNLIPGGFPPTAMPPLTYQL
SEHSYQRTSIQAVSDSFSTVHRPQPLPPSTVHQSTIPSNPDSSAYCLPSTRHGFS
YTDSFVPSPGSPNMPNTIGNLSPNSIRHNLHLKFIKRVONEGTGKSSWWMLNPE
GKSGSKPRRAASMDNNKFAKRSRAAKKASIOSQEGAGDPSGOSFKWSPASPG
SHSNDNDNWTFTPRSSNAATISGRSLPINTQDDLGEGDVHSMYTPPSAAKAST
LPSUSEISNPENMENDLNLSTPLTSLVSTQSSPTMMQQTFCYSFAPNPTSLNS
PSPNYKYTYGQSSMLPQMPIQTLDQNKSSYGGMSQYNCAKGLKELLTSDSPPHN
DINTPVPDGVAPNSRVLGQNMVMPNSVMSYTGQASHNNKMPSPSHTHPGHAQOTS
AVNGRPLPHTVSTWPTSGMRLTQVKTPTQVPLPHPMQMSALGYSVSSVCYGRM
GLLHQELPSLDGMFIERLDCMESIIRNDLMDGDTILDFFDNLVLPNQSPPHSVKTT
THSWVSG"
ORIGIN
    Query Match 100.0%; Score 20; DB 8; Length 3517;
    Best Local Similarity 100.0%; Pred. No. 0.28;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20
    |||||||
Db 2022 GCTTTGGTTGGCAACACAT 2003

RESULT 20
CQ724998/c 4714 bp DNA linear PAT 03-FEB-2004
LOCUS CQ724998
DEFINITION Sequence 10932 from Patent WO02068579.
ACCESSION CQ724998
VERSION CQ724998.1 GI:42285855
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
REFERENCE 1
    Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
    Kits, such as nucleic acid arrays, comprising a majority of
    humanexons or transcripts, for detecting expression and other uses
    thereof
    Patent: WO 02068579-A 10932 06-SEP-2002;
    PE Corporation (NY) (US)
FEATURES
    source
        Location/Qualifiers
            1..4714
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
    Query Match 100.0%; Score 20; DB 6; Length 4714;
    Best Local Similarity 100.0%; Pred. No. 0.27;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20
    |||||||
Db 1297 GCTTTGGTTGGCAACACAT 1278
```

```
RESULT 21
HSM808426/c 4919 bp mRNA linear PRI 30-AUG-2003
LOCUS HSM808426
DEFINITION Homo sapiens mRNA; cDNA DKFZp686B20125 (from clone DKFZp686B20125).
ACCESSION BX648278
VERSION BX648278.1 GI:34367437
KEYWORDS
    Homo sapiens (human)
SOURCE Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
REFERENCE 1 (bases 1 to 4919)
    Bloecher,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Oeanger,A.,
    Fobo,G., Han,M. and Wiemann,S.
    The German Human cDNA Consortium
    Direct Submission
    Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
    Neuherberg, GERMANY
    Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
    Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
    sequenced by GBF (National Research Centre for Biotechnology Ltd.,
    Braunschweig/Germany) within the cDNA sequencing consortium of the
    German Genome Project.
    This clone (DKFZp686B20125) is available at the RZPD in Berlin.
    Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
    Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
    information about the clone and the sequencing project is available
    at http://mips.gsf.de/proj/cDNA/.
FEATURES
    Location/Qualifiers
        1..4919
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp686B20125"
            /tissue type="human uterus"
            /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
            DL10B; sites SfiIA + SfiIB"
            /dev_stage="adult"
            4844
polyA_site
ORIGIN
    Query Match 100.0%; Score 20; DB 8; Length 4919;
    Best Local Similarity 100.0%; Pred. No. 0.27;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20
    |||||||
Db 1508 GCTTTGGTTGGCAACACAT 1489

RESULT 22
MMU252157/c 4945 bp mRNA linear ROD 15-APR-2005
LOCUS MMU252157
DEFINITION Mus musculus mRNA for forkhead protein FKHR (FKhr gene).
ACCESSION AJ252157
VERSION AJ252157.1 GI:7530135
KEYWORDS FKhr gene; forkhead protein FKHR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
    Leenders,H., Benoit,C. and Mathis,D.
    The forkhead FKHR is involved in thymocyte proliferation
    Unpublished
    JOURNAL
REFERENCE 2 (bases 1 to 4945)
    Leenders,H.
    Direct Submission
    Submitted (06-JAN-2000) Leenders H., Alsace, IGBMC, BP 163, 67404
```



```
Qy 1 GCTTTGGTTGGCAACACAT 20
Db 2305 GCTTTGGTTGGCAACACAT 2286

RESULT 26
AF032885/c
LOCUS AF032885 Homo sapiens forkhead protein (FKHR) mRNA, complete cds.
DEFINITION AF032885 Homo sapiens forkhead protein (FKHR) mRNA, complete cds.
ACCESSION AF032885
VERSION AF032885.1 GI:2895491
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 5723)
AUTHORS Anderson,M.J., Viars,C.S., Czekay,S., Cavenee,W.K. and Arden,K.C.
TITLE Cloning and characterization of three human forkhead genes that
comprise an FKHR-like gene subfamily
JOURNAL Genomics 47 (2), 187-199 (1998)
PUBMED 9479491
REFERENCE 2 (bases 1 to 5723)
AUTHORS Anderson,M.J., Viars,C.S., Czekay,S., Cavenee,W.K. and Arden,K.C.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1997) Medicine, Ludwig Institute for Cancer
Research, San Diego Branch, 9500 Gilman Drive, La Jolla, CA
92093-0660, USA
FEATURES
source
Location/Qualifiers
1..5723
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="13"
/map="13q13-q14.1"
1..5723
/gene="FKHR"
386..2353
/gene="FKHR"
/codon_start=1
/product="forkhead protein"
/protein_id="AAC39591.1"
/db_xref="GI:2895492"
/translatlon="MARAPQVETDPDEPLPRSCVTWLPPEFSQNSATSSPAP
SGSAANPDAAAGLPSASAAVSADFMSNLSLLESEDFPQAPQSVAAVAARAAA
TGCGGDFQGEAGCLHPAPPPPPGLSPHPPVPAAGPLAGQPRKSSSSRRNAW
GNLSYADLITKAIESSAEKRLTSLQIYEMVMKSYVYFKDKGSSAGWKNISIRHNS
LHSEFIRVNEGTEGKSWMLNPEGKSGKSPRRRAASMDNNSKFAKSRRAAKKAS
LQSQEGAGSPGQSKWPAKPGSHNDPFDNMTSPRTSSNASTISGRLSIPMTE
QDDLEGDVHSMVYPPSAKMAWTLPSLSISNPNMENLNLNLSSPSTLTVSTQ
SSPCTMQQTCYSPAPNTSLNSPSPNYQKTYGQSMSPLPQMPQTQDNKSSYG
GMSYNCAPGLKELLTSDSPPHNDIMTPVDGVAQNSRVLGQVWGNVSNVSTYG
SQASHKAMPSSHTPEGHAQQTSAVNGRPLPHTVSTMPHTSGMNRLTQVKTQVQVPL
PHPMQALGYSVSCNGYGRMGLLHQELKPSDLDCMFIERLDCMESIIRNDLMD
GDTLDFDNVLPNQSPPHSVKTTTHSWVSG"

Qy 1 GCTTTGGTTGGCAACACAT 20
Db 2305 GCTTTGGTTGGCAACACAT 2286

RESULT 27
AX779990/c
LOCUS AX779990 Homo sapiens forkhead protein (FKHR) mRNA, complete cds.
DEFINITION AX779990 Sequence 2147 from Patent WO03039443.
ACCESSION AX779990
VERSION AX779990.1 GI:32696984

Qy 1 GCTTTGGTTGGCAACACAT 20
Db 2305 GCTTTGGTTGGCAACACAT 2286

RESULT 28
AL133318
LOCUS AL133318 81862 bp DNA linear PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RP11-89L15 on chromosome 13 Contains
the 3' end of the FOXO1A gene for forkhead box O1A
(rhabdomyosarcoma), complete sequence.
ACCESSION AL133318
VERSION AL133318.11 GI:13784997
KEYWORDS HTG; FOXO1A; rhabdomyosarcoma.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 81862)
AUTHORS Beasley,H.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
COMMENT On Apr 24, 2001 this sequence version replaced gi:13374937.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 13, constructed by the Sanger Centre Chromosome 13
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr13
RP11-89L15 is from the library RPC1-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
```





Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Vassiliev,H., Venkataranan,V.S., Viel,R.,  
Vo.A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
Zimmer,A. and Zody,M.

## Direct Submission

Submitted (21-JUL-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

On Jul 21, 2005 this sequence version replaced gi:67003692.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@broad.mit.edu

----- Project Information

Center project name: L32659

Center clone name: 212\_F\_17

-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 17312: contig of 17312 bp in length  
\* 17313 17412: gap of unknown length  
\* 17413 59435: contig of 42023 bp in length  
\* 59436 59535: gap of unknown length  
\* 59536 94648: contig of 35113 bp in length  
\* 94649 94748: gap of unknown length  
\* 94749 149841: contig of 55093 bp in length.

## FEATURES

source

1. 149841  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="3"  
/map="3"  
/clone="RP24-212F17"  
/clone\_lib="RPCI-24 Male Mouse BAC"  
17313..17412  
/estimated\_length=unknown  
59436..59535  
/estimated\_length=unknown  
94649..94748  
/estimated\_length=unknown

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 149841;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20

Db 86068 GCTTTGGTTGGGCAACACAT 86049

## RESULT 30

AC078964

LOCUS

AC078964 155595 bp DNA linear HTG 15-JAN-2001

DEFINITION Homo sapiens chromosome 8 clone RP11-135C3 map 8, WORKING DRAFT

SEQUENCE, 23 unordered pieces.

AC078964

AC078964.2 GI:12229296

VERSION HTGS PHASE1; HTGS\_DRAFT.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

## Hominidae; Homo.

1 (bases 1 to 155595)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 8, clone RP11-135C3

## Unpublished

2 (bases 1 to 155595)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bida,P., Boguslavsky,L.,  
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,  
Choepl,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,  
Darellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,A., Horton,L.,  
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,  
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,  
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,  
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sougnaz,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo.A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.

## Direct Submission

Submitted (13-AUG-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jan 15, 2001 this sequence version replaced gi:9799780.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L10496

Center clone name: L35\_C\_3

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 144191 bases at least Q40

Consensus quality: 150273 bases at least Q30

Consensus quality: 152425 bases at least Q20

Insert size: 160000; agarose-fp

Insert size: 153395; sum-of-contigs

Quality coverage: 3.8 in Q20 bases; agarose-fp

Quality coverage: 4.0 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 23 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1089: contig of 1089 bp in length  
\* 1090 1189: gap of 100 bp  
\* 1190 2528: contig of 1339 bp in length  
\* 2529 2628: gap of 100 bp  
\* 2629 5195: contig of 2567 bp in length  
\* 5196 5295: gap of 100 bp  
\* 5296 6629: contig of 1334 bp in length  
\* 6630 6729: gap of 100 bp  
\* 6730 8429: contig of 1700 bp in length  
\* 8430 8529: gap of 100 bp  
\* 8530 30093: contig of 21564 bp in length  
\* 30094 30193: gap of 100 bp





```

misc_feature 16506 .19528
/note="assembly_fragment"
gap 19529 .19628
/estimated_length=100
misc_feature 19629 .22463
/note="assembly_fragment"
gap 22464 .22563
/estimated_length=100
misc_feature 22564 .62728
/note="assembly_fragment"
gap 62729 .62828
/estimated_length=100
misc_feature 62829 .65273
/note="assembly_fragment"
gap 65274 .85373
/estimated_length=100
misc_feature 65374 .69999
/note="assembly_fragment"
gap 70000 .70099
/estimated_length=100
misc_feature 70100 .73397
/note="assembly_fragment"
gap 73398 .73497
/estimated_length=100
misc_feature 73498 .80679
/note="assembly_fragment"
gap 80680 .80779
/estimated_length=100
misc_feature 80780 .89098
/note="assembly_fragment"
gap 89099 .89198
/estimated_length=100
misc_feature 89199 .96536
/note="assembly_fragment"
gap 96537 .96636
/estimated_length=100
misc_feature 96637 .102440

Query Match 100.0%; Score 20; DB 14; Length 186044;
Best Local Similarity 100.0%; Pred. NO. 0.18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCACACAT 20
|||||
Db 154550 GCTTTGGTTGGGCACACAT 154531

RESULT 32
AC098951/c
LOCUS AC098951
DEFINITION Rattus norvegicus clone CH230-173B15, WORKING DRAFT SEQUENCE, 5
unordered pieces.
ACCESSION AC098951
VERSION AC098951.5 GI:30581568
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 228076)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

```

```

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Georgievski,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoggess,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwokedemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu.L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.B.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajda,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umanik,K.,
Valas,R., Vera,V., Villanasa,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 228076)
Worley,K.C.
Direct Submission
Submitted (07-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 228076)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23100953.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GIJG
Center clone name: CH230-173B15
----- Summary Statistics

```



assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GSHZ  
Center clone name: CH230-216D16  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 238069 bases at least Q40  
Consensus quality: 241686 bases at least Q30  
Consensus quality: 243440 bases at least Q20  
Estimated insert size: 246925; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 95547: contig of 95547 bp in length  
\* 95548 95647: gap of unknown length  
\* 95648 295240: contig of 199593 bp in length  
\* 295241 295340: gap of unknown length  
\* 295341 296872: contig of 1532 bp in length  
\* 296873 296972: gap of unknown length  
\* 296973 303046: contig of 6074 bp in length.

FEATURES

source  
1. 303046  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-216D16"  
1. 1110  
/note="wgs contig"  
95548. .95647  
/estimated\_length=unknown  
95648. .96787  
/note="wgs contig"  
119792. .121768  
/note="wgs contig"  
289420. .291216  
/note="wgs contig"  
295241. .295340  
/estimated\_length=unknown  
296873. .296972  
/estimated\_length=unknown

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 303046;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACAT 20

Db 75187 GCTTTGGTTGGCAACAT 75168

RESULT 34

AE016828.11/c  
WFCOMMENT

Sequence split into 20 fragments LOCUS AE016828 Accession AE016828

Fragment Name Begin End  
AE016828\_00 1 110000  
AE016828\_01 100001 210000  
AE016828\_02 200001 310000  
AE016828\_03 300001 410000  
AE016828\_04 400001 510000  
AE016828\_05 500001 610000  
AE016828\_06 600001 710000  
AE016828\_07 700001 810000  
AE016828\_08 800001 910000  
AE016828\_09 900001 1010000  
AE016828\_10 1000001 1110000  
AE016828\_11 1100001 1210000  
AE016828\_12 1200001 1310000  
AE016828\_13 1300001 1410000  
AE016828\_14 1400001 1510000  
AE016828\_15 1500001 1610000  
AE016828\_16 1600001 1710000  
AE016828\_17 1700001 1810000  
AE016828\_18 1800001 1910000  
AE016828\_19 1900001 1995281

Continuation (12 of 20) of AE016828 from base 1100001 (AE016828 Cxiella burnetii RSA 49

Query Match 95.0%; Score 19; DB 1; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 0.75;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACA 19

Db 66215 GCTTTGGTTGGCAACACA 66197

RESULT 35

AJ834861

LOCUS

DEFINITION

Arabidopsis thaliana T-DNA flanking sequence, left border, clone

074G04.

ACCESSION

AJ834861

VERSION

AJ834861.1

GI:52539067

KEYWORDS

left border; T-DNA flanking sequence.

SOURCE

Arabidopsis thaliana

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1

Branaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F.,

Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,

Lepiniec, L., Caboche, M. and Lecharny, A.

T-DNA integration into the Arabidopsis genome depends on sequences

of pre-insertion sites

EMBO Rep. 3 (12), 1152-1157 (2002)

PUBMED

12446565

REFERENCE

2

(bases 1 to 299)

AUTHORS

Balzerque, S.

TITLE

Direct Submision

JOURNAL

Submitted (21-SEP-2004) Balzerque S., UMRGV, INRA/CNRS, 2 rue

Gaston Cremieux, 91057 Evry cedex, FRANCE

PCR was performed on DNA from transformants of Arabidopsis thaliana

plants from INRA (Versailles). The DNA fragment (s) resulting from

the PCR were directly sequenced from the left or the right border

to determine the genomic sequence flanking the insertion. T-DNA

derived sequences were removed. Information to order the

corresponding mutant line and a link to a database providing a

graphical display of the insertion site are available at

<http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has

been generated in the framework of the French plant genomics

program 'Genoplante' (<http://www.genoplante.com> and

<http://genoplante-info.infobiogen.fr>).

Location/Qualifiers

FEATURES





VNADLPTRPAAPFOFTTVDLFGPVLKDDKREVTLTGTWGVFSCWACRAIHLDL  
VNSVSESLMAYQRTAIRGHPSKLSDPGTNFICAKPVLQDLQFLESQNKALAE  
YAVSGTEWRQIHPDSDPHRGGAAGAAVRKARALQALDKNTLSYFQVLFATA  
NLANERPIEARVOSKEDCIRYVSPNSLLGRASNGDFNSPDFESYSYRLQVIOCEV  
NNFWKSVQLVGNPFIIRNKWHTLRNVSEGDIVWCDQNSLRGQVKKVKNPDARGI  
ISVDVKKVQSHCIPMKFHVNSVCKTAKDKLSHTVLHRDVRRLVLLPVEGKKG  
ISTVEKN"  
variation  
6418  
/gene="gag-pol"  
/note="results in premature stop in Fugu Sequencing  
Project Scaffold M004397"  
/replaces="t"  
6882..6891  
/note="polypurine tract"  
6892..7613  
/note="3' LTR"  
7614..7618  
/note="integration site duplication"  
/rpt\_type=direct  
ORIGIN  
Query Match 85.0%; Score 17; DB 5; Length 7618;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 TTTGGTTGGCAACACA 19  
|||||  
Db 4168 TTTGGTTGGCAACACA 4152  
RESULT 38  
FRIGVHL  
LOCUS  
DEFINITION Fugu rubripes immunoglobulin heavy chain gene cluster, partial  
sequence.  
ACCESSION AF108421  
VERSION AF108421.1 GI:4633112  
KEYWORDS  
SEGMENT  
SOURCE 1 of 3  
ORGANISM Takifugu rubripes  
Takifugu rubripes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae; Takifugu.  
REFERENCE 1 (bases 1 to 40406)  
AUTHORS Peixoto,B.R. and Brenner,S.  
TITLE Characterization of approximately 50 kb of the immunoglobulin VH  
locus of the Japanese pufferfish, Fugu rubripes  
JOURNAL Immunogenetics 51 (6), 443-451 (2000)  
PUBMED 10866111  
REFERENCE 2 (bases 1 to 40406)  
AUTHORS Peixoto,B.R. and Brenner,S.  
TITLE Direct Submission  
JOURNAL Submitted (23-NOV-1998) The Molecular Sciences Institute, 2168  
Shattuck Ave. 2nd Floor, Berkeley, CA 94704, USA  
FEATURES  
source  
Location/Qualifiers  
1..40406  
/organism="Takifugu rubripes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:31033"  
/germline  
complement(864..4244)  
/note="similar to complement decay accelerating factor 1  
gene"  
misc\_feature  
complement(9009..12919)  
/note="similar to calmodulin gene"  
Join(16193..16241,16326..16628)  
/note="immunoglobulin heavy chain V segment 1 (F-Vh1)"  
18282..22726  
/note="putative retrotransposon"  
Join(24516..24555,24645..24937)  
/note="immunoglobulin heavy chain V segment 2 (F-Vh2)"

V\_segment  
Join(25958..25906,25992..26293)  
/note="immunoglobulin heavy chain V segment 3 (F-Vh3)"  
Join(27181..27223,27317..27612)  
/note="immunoglobulin heavy chain V segment 4 (F-Vh4)"  
Join(28563..28611,28697..28998)  
/note="immunoglobulin heavy chain V segment 5 (F-Vh5)"  
CDS  
complement(33520..39177)  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAD26627.1"  
/db\_xref="GI:4633116"  
/translation="MAVATGKTVSELRKDRVTAKRLTKVKNITSRNHEDNNEBELK  
DILSKLSEAEKLEANDDLQAGIADQAEEDSGEAVLDQVRLDLEKTSQDCRRR  
LAERENYKQSLWNSFQAADLSIAVTTETVVDYDPAITQMRPEDPGVGVHIVISQLOELF  
KRAEENYKTSWIPGHNREIQDQRELEFIPALITRKADFIKVTYKDDMPQAAA  
TVNYSIPALRLKPTALPRFNGCQGFQWRKQWEALQSQEPTSPVRRNQPKI VELL  
EKIKDLRLSYSTADEIFRVLESRYGNKTFALIEVELORTSPVRRNQPKI VELL  
RTVEKALQDLCELGDTGAIKNPLVTKSIESKLPDSLKEMLIYATDTNNAVQPEKFD  
CLLAPLKQEPITYEQLELLLEBPVRKEARAEHLHARTKTVKSDGQACVCGDVKH  
KKLYFCQFRSSKLEDDKQAAVRRIRACPKLEIHGDELCKCPGLCKGQRCIDGRAR  
HHYLCPYAKESKNDPRDRHSKEGPAKAFTEQDFSRSLSPELAKQCDVFNSEA  
SKTFNAVNSHSLAESGLQELPVMLELVATANAGRLGLTILIASDTNYITHSAA  
RLNLRSEBITLVHGVGMEVCVETKRYLLRIRVKTSGTLKSHQVCGVGLDRIARVH  
QSVTAGOLQRFPRIAQAELEPKKEVDLLISHREGELAPQVQVAGDILVWGPGLKT  
VGGAPHLPEKLVLSAHASRTHFAKSMRTAAVRYVMAGTPPKPLNSHYIKTNSAH  
QEFLLKMRWDSIGTACEPKCGYRCNGCPGCKEMLTAEERELEVRAGLTHVMSDDH  
STEPHHAKYPMWDDQLQPNRNNAQATFLMRKELAKEPNKVAAYAAQVHEMLSRG  
AAAKLSDSITAEWTGPYIYISHLIAPNPHSVTPVRLVWNSQRFKGSALDMLLKG  
DVLNQIRAVLLRFRSGVAAIGDVRKMTNSVLEDRVHLHRLWRNSAEERIOEVAV  
TVNIGDPAGCIAQLAMETASLPQESHLTAREVLRHNAVVDLILSHNDLKQLQA  
TTKGVEDILKAGGPAKPFSGSGRGDHDIDRQETVQDAPFLVNQMSDDNDKALGL  
GYIVKBEKLVWVAINFSSKKGMHLGDLQPPQVRFCTPSPTRRELLSQVGLYDP  
VGLVAPVKQGAIIIVRRFAQSEHQHPADTKWALSAELREDAILFEECVRLRDQ  
KFDRLATPSNFIKGPWALTFGDSEHSQVAVYLRWETDESPILALVEAKAKSLDQ  
KGEAVKCGAVIASRLKYPQSHSPIQVDRWFHVDSTVLGAIORECYGQTFPS  
NRIGIOTNSRAEDMMWVPGPNVADMITRGNASPELSEHSIQWRGPDLKPIDTPW  
VKSADTATNARENISKQKAFVDAITRSMANDPKQWTSKASRRSPASGEIKHL  
VDARFSTLAKPIKTVALVWRAAKNLFSLKDVENKYSVSKWEAVSSKGLISVKE  
DALRDI FLATQCGQFQDSTNRLLVYQDSTGLLAVSGRVSFNDNRVPLLPYDA  
WVSTLMAREAHGSHDVGAGTLLKMRCKAWIKGRVRAQVVDGCLCKRIARRCQ  
VMADLPPTPAAPFOFTTVDLFGPVLKDDKREVTLTGTWGVFSCWACRAIHLDL  
VNSVSESLMAYQRTAIRGHPSKLSDPGTNFICAKPVLQDLQFLESQNKALAE  
YAVSGTEWRQIHPDSDPHRGGAAGAAVRKARALQALDKNTLSYFQVLFATA  
NLANERPIEARVOSKEDCIRYVSPNSLLGRASNGDFNSPDF"  
ORIGIN  
Query Match 85.0%; Score 17; DB 5; Length 40406;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 TTTGGTTGGCAACACA 19  
|||||  
Db 35772 TTTGGTTGGCAACACA 35788  
RESULT 39  
AB024035/c  
LOCUS  
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone:MHM17.  
ACCESSION AB024035 BA000015  
VERSION AB024035.1 GI:4519194  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1  
Sato,S., Nakamura,Y., Kaneko,T., Katoh,T., Asamizu,E., Kotani,H.  
and Tabata,S.  
STRUCTURAL ANALYSIS OF ARABIDOPSIS THALIANA CHROMOSOME 5. X.  
Sequence features of the regions of 3,076,755 bp covered by sixty  
Pl and TAC clones

JOURNAL PUBMED REFERENCE AUTHORS TITLE JOURNAL	DNA Res. 7 (1), 31-63 (2000) 10718197 2 (bases 1 to 78423) Nakamura.Y. Direct Submission Submitted (24-FEB-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan [E-mail: ynakamura@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934] Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see <a href="http://www.kazusa.or.jp/kaos/cgi-bin/aggd_graph.cgi?c=MHM17">http://www.kazusa.or.jp/kaos/cgi-bin/aggd_graph.cgi?c=MHM17</a> Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <a href="http://compbio.ornl.gov/Grail-1.3/">http://compbio.ornl.gov/Grail-1.3/</a> ), GENSCAN (Chris Burge, MIT, <a href="http://CCR-081.mit.edu/GENSCAN.html">http://CCR-081.mit.edu/GENSCAN.html</a> ), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <a href="http://www.cbs.dtu.dk/services/NetGene2/">http://www.cbs.dtu.dk/services/NetGene2/</a> ) and SplicePredictor (Volker Brendel, Stanford University, <a href="http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi">http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi</a> ). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <a href="http://genome.wustl.edu/eddy/tRNAscan-SE/">http://genome.wustl.edu/eddy/tRNAscan-SE/</a> ). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MP10 and the 3' clone is MUL3. Location/Qualifiers 1. .78423 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="5" /clone="MHM17" /clone_lib="Mitsui P1" /ecotype="Columbia" complement (1. .30) /note="CDS is reported in Acc# AB020747 contains similarity to protein kinase gene_id:MP10.5" /number=1 /evidence=not experimental complement (join(4221. .4349,4616. .4761,4851. .4944, 5022. .5307,5397. .5821,6051. .6208,6305. .6432,6529. .6638, 6709. .6982,7070. .7141)) /note="unannotated protein product; dbj BAA91947.1 gene_id:MHM17.1 similar to unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAA97020.1" /db_xref="GI:8777430" /translation="MAPRIILCGDPLGRNLQFLKRVQSVKSAGPFDALICVQGFPPD SPEILDFLDYVEGRAQVPIPTPTGDTGYGVVAPKILSTTSKAEQKMGDLGVCHN LFLWRGSGKSLVGLSVLQSGSSQFGKYSDDDVALRALAEPPGCTIWEPA GVTRAAVSDIPVIGISDSSCDSTVSELVMEKVRVHIAGMGVFAVEPYLAESTH VTRFGLAQVGNKKQKFLHALSPTPTSTWSPALSAKPKPTLWPNLQDGAESKK RPNDSDSQWRYDVPKRSQSGGKLCFKFVCSGSCPRGSDCHPHQNAEBOCR RGWCLDIIIGKECKEPCSYRHEFODESIIQRKPSNANRKECWFCISPSVESH LIYVSGSFYCALPFGSLGVDHILIIPIEHLNPTLVLSPEVSELSRYQNGRLCYKS QGNDVFFELVSKRVSHANLQVVPSPSRALLPNI FSLAEKLGPKLVTKFNDS TD GRXYLQKEYNAALGLFVELPDGTVLSHTLEENEVPPAQFGREVLGLLXIPDRADR NCKLSQEEAKLAEFPKQGFDPCCQ" complement (7467. .7643,7741. .8080,8177. .8230,8376. .8479) /note="unannotated protein product; gene_id:MHM17.2 pir  C71422 similar to unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAA97021.1" /db_xref="GI:8777431"
FEATURES source	translation="MADSVLKBYNCGRPEKIPKLDKACEGSKSWKHLKLGNDVDEEDY LROYCLPHYFHKSEGTVDWEKYDMFHIRPLENSPPIISDRTNADVIDVTLTFAIE KNEAGSKLVFVEHYVSANFKFANGLLCMLTFWATMDASSPQSQIYQVLELRRGKQF EIPFVRVKDEEMDDVEVPFSPMPYDDYDKPVPVFRAPEDGVFPVFDRTGALYD LYRSGL" join(9471. .9653,9736. .10236) /note="unannotated protein product; gene_id:MHM17.3 pir  C71422 similar to unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAA97022.1" /db_xref="GI:8777432" /translation="MADSVLKBYNRRPEKIPKLDCEGSESSSKORVVKWGTAEDD EYLROYLLPHYQKQCGFSIKWEQFDYNFRSRMDKSPNEISERKSNVILIRMTLT ALDKHNEAGTKLVFVEHYEANYQLTREGLCMLTFWATMDASSPQSQIKYQAHVWRG TGAFHTIFIRUKPTDREIEAVEVFPSPMLYELDKPPIVFRPGEDEALPGVPVFRN TGAGLDPDW" complement (join(12420. .13157,13254. .13394,13492. .13542, 13628. .13759,13862. .13927,14005. .14747,14880. .15036)) /note="unannotated protein product; gene_id:MHM17.4 unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAA97023.1" /db_xref="GI:8777433" /translation="MESSVAPFFHRRTHLPNRYRSLLYHFCSSFDREPOISLATPAV LQBLVRLTEIAQSPGETCEPPENLSTITESKLGVSQSGEKEVETISQSKSLMGLDI CGLDLDQASVPSKRHTDFFDSFELMINETQDSVPESCYNLFALDYNVDYDIQVNLK PNIATQVQDVPVESEKAEVPSVESNEVISGVEACNLTQVOREMDEKPDVNSPV LVDVSRIYVGGDDVESEIGDDNDMDVEDDETVERHEEYQVSDQGTNSHTSHKS FGVEVMVNDQAKKIQDTFSNBAKDPGTSIKKRSAPSKADAKAKAKARIKKAQERI ALGVKLLKVPAPKPKLYKCHRYLKGRCHEGDKCFSDHTTIPETKSCPCCYFATOS CMKDDCPDHDLSKYPCNNFITKFCYRGDSCLPFKHGTGPQASDTPSPNAVTSSTK ITAASFSPQTKQSVRDAIAKLPAPQARVSSVAFLKPSHSNQRNSDASSKINE HTVPQVPPLPKPSVAPKGMFSFLDKTQEDTVTKASSAKNPNTDSDSQTLKQSQG SFPLPGPKKISPLSPASBEQKTLNRPQKPASSKNLKTTPSSHIOSSLISAMKLA E FESAKYRGNNDPTEAVNKSNTVDTAVTRNSGNISSKILEFLSSFSHGKN" join(15350. .15482,15570. .15615,16443. .16671) /note="gene_id:MHM17.5" /codon_start=1 /evidence=not experimental /product="30S ribosomal protein S16" /protein_id="BAA97024.1" /db_xref="GI:8777434" /translation="MVVRFLSPFGCKNPPFFRVMAADSRSPEDGKHLEVLGVFNPLP QDQGRKMLGKFDRIKYLWSVGAQPSDQVQRLLPFRSGLLPPPPVMVAMRGKGARDTRP VDPMTGRYVDAENKTVNANDNQPKBEDTEAKSA" join(17089. .17134,17217. .17260,17389. .17484,17694. .17786, 17919. .18006,18095. .18156,18245. .18429,18562. .18739, 18827. .18870,19017. .19218,19684. .19770) /note="gene_id:MHM17.6" /codon_start=1 /evidence=not experimental /product="nucleosome assembly protein" /protein_id="BAA97025.1" /db_xref="GI:8777435" /translation="MSNDKDSNVSDLTSLAKDEDRAGLVNALKNKLONLAGOHSDDL ENUTPKRRVRVRLIRIQGHDEIFKFRERAALEAKYQKLQPLYNKRYEIVNGAT EYVGAEPDAMDQGDGTAEKGVFPWLTALKNDVISEBITERDEGALILYLDIKW CKIEBKPKLFEFFDQNFYFKNTLLTKAYHIMDEDEPLELEKAIPTGTEIDWYPGKCLTQ KILKPKPKGAKNAKTIKTDECESEFFNFNPQVDDDDIDDEBAEELQNLMEQDY DIGSTTRREKII PHAVSWFTGEALGEFEFIDDDDDIDDEDDEDEDEDEDEDE DEEEYSYTKKKEFIILQKGRPOVTDQOGERPPECQK" complement (join(23174. .23239,23666. .24049,24174. .24863, 25283. .25388,25491. .25645)) /note="unannotated protein product; gb AAC78547.1 gene_id:MHM17.7 similar to unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BAA97026.1" /db_xref="GI:8777436"
exon	
CDS	
CDS	
CDS	

```
/translation="MMHLTSCSYLISMDGYNEASERBSSSSSSGSLARSLFHEVRO
SVPLQNGHVPMSAFNNLPPYIEIPQESQRLAFNDTQRLFYQMKIEASLREWFEDF
NRKSPANSDIYRPHYPSSSSSISPNNISILLPLLPKPTTITAVNVPVLP
LAPINIMHPQHEPLFRNRQREBEAMTQAILAVLTGPSPSPSTSPQKGRATAPKR
YYSMSIDRGAFLPVRKQSMWTRAMSFYNNLRINQRRFTRENATTHGEGSGGGG
GRYTSQATQOHMISLRKREKLNESFOALRSLLPGTKKDKASVLSIAREQLSSL
QGEISKLENNREAKLAGREIEIENDLRPERFNRIRHIESRSRRTDLRLVLR
GDIIRVDLMIRLLEPLKQINNVLVSIEARTLAREAGDTSIVLISLRLKEGEWDE
SAFQZARRRVADLAH"
CDS
30529..30656,32107..32716)
/note="gene id:MHM17.8"
/codon_start=1
/evidence=not experimental
/product="cytokinin oxidase"
/protein_id="BAA97027.1"
/db_xref="GI:8777437"
/translation="MASYNLRQSVRLIAITVIIITLSTPIITNTSPQWNLISHNEF
AGKLTSSSSVESAAFDGHVTKIPSAVLIPSSVEDITDLKLSFDSQLPPLAARG
HGHSHRGQASAKGVVNMRSVMNRDGIKVSRCTCLYVDVDAWLMIEVLKLELGL
TPVSHVDLYLTGVGTLSNGISGQTFRGPQITNVLEMDVITGGEIATCSKWNSD
LFFAVLGGGQGIITRARIKLEVAPKRAKMLRFLYIDFSEFRQERVIKTDGVDF
LEGSTMDHPGPDNRSTYPPSDHLRIASVMKRVIVCYLEVVKYIDETQYVNEE
MEELSDSLNHRVGFYKDVYMDFLNRVTGELNKLKGQWDVPHPLNLFVPTQI
SKFDGVFKGILRNITSGPVLVPMNRKNWDRMSAIPREEDVFYAVGLFSAGFD
NWEAFQENMEILKECEDANMGVIOQLPHSSQEGWVRHFGFRWNIFFERYKYDPMK
ILSPQNIQKINKNS"
tRNA
complement(39239..39315)
/product="tRNA-Val"
/note="codon recognized: GUG; gene_id:MHM17.9"
/evidence=not experimental
complement(join(41045..41748,41932..42184))
/note="unnamed protein product; gene_id:MHM17.10
pir||T04268
/codon_start=1
/evidence=not experimental
/protein_id="BAA97028.2"

Query Match      85.0%; Score 17; DB 15; Length 78423;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACA 17
Db 70517 GCTTTGGTTGGCAACA 70501

RESULT 40
CP000084_08/c
WPCOMMENT
Sequence split into 13 fragments LOCUS CP000084 Accession CP000084
Fragment Name      Begin      End
CP000084_00        1         110000
CP000084_01       100001     210000
CP000084_02       200001     310000
CP000084_03       300001     410000
CP000084_04       400001     510000
CP000084_05       500001     610000
CP000084_06       600001     710000
CP000084_07       700001     810000
CP000084_08       800001     910000
CP000084_09       900001    1010000
CP000084_10      1000001    1110000
CP000084_11      1100001    1210000
CP000084_12      1200001    1308759
Continuation (9 of 13) of CP000084 from base 800001 (CP000084 Candidatus Pelagibacter ub

Query Match      85.0%; Score 17; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACA 17
```

```
Db 94372 GCTTTGGTTGGCAACA 94356
|||||
RESULT 41
AY144116/c
LOCUS
DEFINITION
AY144116
47740 bp DNA linear BCT 17-JUL-2003
Photorhabdus luminescens strain W14 type III secretion system,
partial sequence; Orf39 (orf39), Orf40 (orf40), Orf41 (orf41), and
Orf42 (orf42) genes, complete cds; Orf43 (orf43) gene, partial cds.
ACCESSION
AY144116
VERSION
AY144116.1 GI:27550042
KEYWORDS
ORGANISM
Photorhabdus luminescens
Photorhabdus luminescens
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Photorhabdus.
REFERENCE
1 (bases 1 to 47740)
Waterfield,N.R., Daborn,P.J. and ffrench-Constant,R.H.
Genomic islands in Photorhabdus
Trends Microbiol. 10 (12), 541-545 (2002)
12564983
REFERENCE
2 (bases 1 to 47740)
Waterfield,N.R. and ffrench-Constant,R.H.
Direct Submission
Submitted (22-AUG-2002) Biology and Biochemistry, University of
Bath, Claverton Down, Bath BA2 7AY, UK
FEATURES
Location/Qualifiers
source
1..47740
/organism="Photorhabdus luminescens"
/mol_type="genomic DNA"
/strain="W14"
/db_xref="taxon:29488"
<1..41116
misc_feature
/note="type III secretion system; pathogenicity island"
complement(<1..113)
CDS
/note="similar to acetate CoA-transferase, beta subunit
(atoA)"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAO18069.1"
/db_xref="GI:27550086"
/translation="MNAQQVIAHRAALELQDGVNIGIGIPTQVNYLPD"
complement(116..775)
CDS
/note="similar to acetoacetyl-CoA transferase alpha
subunit"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAO18070.1"
/db_xref="GI:27550087"
/translation="MKNKLSTQQAHEHYDGMTIMAGFGMGVGPPLHITALLNSGV
KDLTIANDTAGIDVIGPLVNNRVKKVITSHGTNPETGRKMLAGEIEVLPLGT
LIERIRCGAGLGSGFLPTVGTGVVBEGKQKITVNGDYDLELPLRADLALIQHLAD
YHGNLTQLSARNPNFVIALAADIETAEPRLEVEGEITPCVITPGALIDYIVYPRG
N"
complement(765..1046)
CDS
/note="similar to response regulator of ato, ornithine
decarboxylase antizyme (sensor ATOS), atoC"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAO18071.1"
/db_xref="GI:27550088"
/translation="MKQEVNQRHRSADNYGWDRLTNSPKQMALCCDTVKVGMPLT
LQAKLYRLNHLISPLGRSEDIALLAYSLPYGHSPMSNISIYGNQYEK"
complement(1269..2351)
CDS
/note="similar to hemolysin from Fusobacterium nucleatum
clinical isolate found in GenBank Accession Number
AF525507"
/codon_start=1
/transl_table=11
```

/product="unknown"  
/protein\_id="AAO18072.1"  
/db\_xref="GI:27550089"  
/translation="MSKEFIWVPEDQSYGHQSQITPSSFDKNI GFTPKMDAGE  
NTLNTDNTNSIKAHDIHPRPSPVELKEQYIVHKAENTDKITGETINISSGNLI  
ISGQKXPVPHLMSQITDRIYIKLQSSSTLAIKANTVRSKPNKTLKPEESAVAI  
SGLSHLVYAESEVQENEVIGKNIISLECFDSTESSKAMKSHLVNIYNSNITLQDN  
AQLISSLINRITDLOEQQLFTVNTFTLKAGATLLNLSNPGTDFPLDIHREDYDK  
GVFNFMAGKENTGKVIDVAPKANAYAGLNTMLRKNFIAINGAVETGDMQKYPDS  
YKTRNDNQVGTIIISLRNPHLKS"  
3583..4038  
/note="hypothetical Orf"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAO18073.1"  
/db\_xref="GI:27550090"  
/translation="MISKTERIAELEKIIQSLSSTSYSVSNFVGNDDVHLTGSSLT  
RTYTMPMETVWGIYASVPIIGPNTDIIKIKPLDFHVMTKYWNVGKILWFOEARAIN  
GVIKYCVGDNFVNLASRAOGNNSGGSKNLKVVCRVPGDLNTAQMIILY"  
complement(4645..5766)  
/note="hypothetical Orf"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAO18074.1"  
/db\_xref="GI:27550091"  
/translation="MFNYKNIHPYEGKLMSTVEFWDPNNSQVYTKDYGTGSIILST  
QSDSFDGLGANTMAESETLITQNDSDNIPYVFPFRNNDGTMSDGLDINVS  
AGTLIVYRLANDQNHVDSHTVAYLGCSSANFILKNSGSLSIKPGTFMFDYV  
SNKPKLIMNSNFEELKOTTKIEDDVPARIFLASELSLESSKLTRESNLYLDGN  
FNYCNISIQKSTVTLNNGIMOKNDIEKNKTFKLAGSPLNLESFDGINFLYLD  
NVYTPEGFLFITTEGNEKGMIEFSGQPNPNKINIKIPDFKGLIAINNKIIEKEK  
HEDYFNTGYKPDQLNDKHEMIITISLKQDQYQSYQYVNV"  
6595..6900  
/note="similar to Erwinia chrysanthemi protease inhibitor  
inh"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAO18075.1"  
/db\_xref="GI:27550092"  
/translation="MKKFMGLVLLGIASANAGADKSIETLDQECIAQWYSSAS  
QSCDLSIWLKNDTCAITACRAVHGVVENGTRFFPYQVSVIANCGGLTKDGC"  
8293..8508  
/note="similar to Yersinia enterocolitica major cold shock  
protein CspA2"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAO18076.1"  
/db\_xref="GI:27550093"  
/translation="MSNMTGSKVWFNNDKGFGFISPKDQKDVHYSAIQSNFXT  
LEEGQVYSFSIENGAKGPSAVNVIVLS"  
complement(8731..9123)  
/note="similar to YopT chaperone SycT from Yersinia  
enterocolitica"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAO18077.1"  
/db\_xref="GI:27550094"  
/translation="WTHDYHQLIHALPSKLGVPASSSTDNIYTLVDRFQVLIGCVH  
DNWLQFLCELGAEISLHANNLFGQHPAHVQDGLDQGTILMSQRPLAEVDVAELQAWLE  
RFIDDIQRLLTPNTTQPKNLPPLMRV"  
complement(9120..10055)  
/note="similar to YopT from Yersinia enterocolitica"  
/codon\_start=1  
/transl\_table=11  
/product="putative type III secreted effector Lopt"  
/protein\_id="AAO18078.1"

/db\_xref="GI:27550095"  
/translation="MSPISGINHORIFPTQOLDTTAVNQPGELSGKSLKVSSRPGL  
LERLSATIQNLPHSMLDQLTDDGKNQESSEPSMIKDRIVHFAVSTKLGSVRDS  
ASHKGVBVYKFAQTGKAFLDQIMKHQDTSGVCESIAHWISAHAGESVFNQLYVG  
QKQGFHIDSLVSIKQLQMDQPYAEQSQITSEWLRENGIQKPSPLEVSGETSGKT  
KDLNALIDTDGSGGYKISFEGWAGHTVAAVYDDQKGVTFDPDFPFDPKTS  
FNNFSDQFMSKMYNKEIGLQNFVYSNPEKTR"  
complement(10231..10848)  
/note="similar to Pseudomonas aeruginosa PscL type III  
export protein"  
/codon\_start=1  
/transl\_table=11  
/product="LscL"  
/protein\_id="AAO18079.1"  
/db\_xref="GI:27550096"  
/translation="MLPFIKITTEHLQAPLQILRKADYQTCLSAQSLLDAARIQAG  
EIERDAQVEYQKELGAGIDAAEAQANXIHQTLQCCQYRQVEQQMSNVLQA  
VRKILKNYDOVSLTLQIVREALSLVSNQKQVILRVNPEQATVREQISRVHKDFPEIG  
YLEITADERLDQGCILETEVGIIDASLDSQLEAVMSAINNQMN"  
complement(10827..11453)  
/gene="lscK"  
/complement(10827..11453)  
/genes="lscK"  
/note="similar to Pseudomonas aeruginosa PscK type III  
export protein; similar to Pseudomonas aeruginosa PscK  
type III export protein"  
/codon\_start=1  
/transl\_table=11  
/product="LscK"  
/protein\_id="AAO18026.1"  
/db\_xref="GI:27550043"  
/translation="MATVLTPOYFRFCPSYIHSDDLPSBWLAVLSLPEWNSPRLN  
GLLTPDLANDVYELPTGLGNLALLQSCLEQLLTLGALLHGOAIBHCLLATSLRL  
HLSLGEKPGFCLKYFDIILIGNWPTQWRSILPAINVNYFRFSLQPLWTAMEPLSID  
FARKLSURLPPYENLAAWPVSAERPLAQLCLKLAKQVNTCYHLLK"  
complement(11454..12191)  
/gene="lscJ"  
/complement(11454..12191)  
/genes="lscJ"  
/note="similar to Pseudomonas aeruginosa PscJ type III  
export protein"  
/codon\_start=1  
/transl\_table=11  
/product="LscJ"  
/protein\_id="AAO18027.1"  
/db\_xref="GI:27550044"  
/translation="MKKPYIICVMLAVMLTGTCKIELYTGVSQKEGEMLALLREAGI  
SSDKPDKDGNIKLLVEESDAQAVEYLRKKGYPRENFSLLQDVQKGLISSPIER  
ARLNFKAQAIARTLSEIDGLVARVHVVLPERQDKLGGKLSPASSSVFIKHADVQ  
DTYIPIQKLVNNSIEGLSYDRISVVLVPAAGVRQVPLAPRYSTLPSIQVTEESQGR  
IGLLVLLALLFVSNLAQFLMHSRMQ"  
complement(12198..12536)  
/gene="lscI"  
/complement(12198..12536)  
/genes="lscI"  
/note="similar to Pseudomonas aeruginosa PscI type III  
export protein"  
/codon\_start=1  
/transl\_table=11  
/product="LscI"

Query Match 80.0%; Score 16; DB 1; Length 47740;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2 CTTTGGTTGGGCAACA 17  
|||||  
Db 30085 CTTTGGTTGGGCAACA 30070  
  
RESULT 42  
AC095920\_09/c  
WPCOMMENT

Sequence split into 10 fragments LOCUS AC095920 Accession AC095920

Fragment Name	Begin	End
AC095920_00	1	110000
AC095920_01	100001	210000
AC095920_02	200001	310000
AC095920_03	300001	410000
AC095920_04	400001	510000
AC095920_05	500001	610000
AC095920_06	600001	710000
AC095920_07	700001	810000
AC095920_08	800001	910000
AC095920_09	900001	1005083

Continuation (10 of 10) of AC095920 from base 900001 (AC095920 Rattus norvegicus clone C

Query Match 80.0%; Score 16; DB 14; Length 105083;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 5 TGGTTGGGCACACAT 20  
|||||  
Db 49873 TGGTTGGGCACACAT 49858

RESULT 43

AP008214\_026

WPCOMMENT

Sequence split into 285 fragments LOCUS AP008214 Accession AP008214

Fragment Name	Begin	End
AP008214_000	1	110000
AP008214_001	100001	210000
AP008214_002	200001	310000
AP008214_003	300001	410000
AP008214_004	400001	510000
AP008214_005	500001	610000
AP008214_006	600001	710000
AP008214_007	700001	810000
AP008214_008	800001	910000
AP008214_009	900001	1010000
AP008214_010	1000001	1100000
AP008214_011	1100001	1210000
AP008214_012	1200001	1310000
AP008214_013	1300001	1410000
AP008214_014	1400001	1510000
AP008214_015	1500001	1610000
AP008214_016	1600001	1710000
AP008214_017	1700001	1810000
AP008214_018	1800001	1910000
AP008214_019	1900001	2010000
AP008214_020	2000001	2110000
AP008214_021	2100001	2210000
AP008214_022	2200001	2310000
AP008214_023	2300001	2410000
AP008214_024	2400001	2510000
AP008214_025	2500001	2610000
AP008214_026	2600001	2710000
AP008214_027	2700001	2810000
AP008214_028	2800001	2910000
AP008214_029	2900001	3010000
AP008214_030	3000001	3110000
AP008214_031	3100001	3210000
AP008214_032	3200001	3310000
AP008214_033	3300001	3410000
AP008214_034	3400001	3510000
AP008214_035	3500001	3610000
AP008214_036	3600001	3710000
AP008214_037	3700001	3810000
AP008214_038	3800001	3910000
AP008214_039	3900001	4010000
AP008214_040	4000001	4110000
AP008214_041	4100001	4210000
AP008214_042	4200001	4310000
AP008214_043	4300001	4410000
AP008214_044	4400001	4510000

AP008214_045	4500001	4610000
AP008214_046	4600001	4710000
AP008214_047	4700001	4810000
AP008214_048	4800001	4910000
AP008214_049	4900001	5010000
AP008214_050	5000001	5110000
AP008214_051	5100001	5210000
AP008214_052	5200001	5310000
AP008214_053	5300001	5410000
AP008214_054	5400001	5510000
AP008214_055	5500001	5610000
AP008214_056	5600001	5710000
AP008214_057	5700001	5810000
AP008214_058	5800001	5910000
AP008214_059	5900001	6010000
AP008214_060	6000001	6110000
AP008214_061	6100001	6210000
AP008214_062	6200001	6310000
AP008214_063	6300001	6410000
AP008214_064	6400001	6510000
AP008214_065	6500001	6610000
AP008214_066	6600001	6710000
AP008214_067	6700001	6810000
AP008214_068	6800001	6910000
AP008214_069	6900001	7010000
AP008214_070	7000001	7110000
AP008214_071	7100001	7210000
AP008214_072	7200001	7310000
AP008214_073	7300001	7410000
AP008214_074	7400001	7510000
AP008214_075	7500001	7610000
AP008214_076	7600001	7710000
AP008214_077	7700001	7810000
AP008214_078	7800001	7910000
AP008214_079	7900001	8010000
AP008214_080	8000001	8110000
AP008214_081	8100001	8210000
AP008214_082	8200001	8310000
AP008214_083	8300001	8410000
AP008214_084	8400001	8510000
AP008214_085	8500001	8610000
AP008214_086	8600001	8710000
AP008214_087	8700001	8810000
AP008214_088	8800001	8910000
AP008214_089	8900001	9010000
AP008214_090	9000001	9110000
AP008214_091	9100001	9210000
AP008214_092	9200001	9310000
AP008214_093	9300001	9410000
AP008214_094	9400001	9510000
AP008214_095	9500001	9610000
AP008214_096	9600001	9710000
AP008214_097	9700001	9810000
AP008214_098	9800001	9910000
AP008214_099	9900001	10010000
AP008214_100	10000001	10110000
AP008214_101	10100001	10210000
AP008214_102	10200001	10310000
AP008214_103	10300001	10410000
AP008214_104	10400001	10510000
AP008214_105	10500001	10610000
AP008214_106	10600001	10710000
AP008214_107	10700001	10810000
AP008214_108	10800001	10910000
AP008214_109	10900001	11010000
AP008214_110	11000001	11110000
AP008214_111	11100001	11210000
AP008214_112	11200001	11310000
AP008214_113	11300001	11410000
AP008214_114	11400001	11510000
AP008214_115	11500001	11610000
AP008214_116	11600001	11710000
AP008214_117	11700001	11810000

```
AP008214_118 11800001 11910000
AP008214_119 11900001 12010000
AP008214_120 12000001 12110000
AP008214_121 12100001 12210000
AP008214_122 12200001 12310000
AP008214_123 12300001 12410000
AP008214_124 12400001 12510000
AP008214_125 12500001 12610000
AP008214_126 12600001 12710000
AP008214_127 12700001 12810000
AP008214_128 12800001 12910000
AP008214_129 12900001 13010000
AP008214_130 13000001 13110000
AP008214_131 13100001 13210000
AP008214_132 13200001 13310000
AP008214_133 13300001 13410000
AP008214_134 13400001 13510000
AP008214_135 13500001 13610000
AP008214_136 13600001 13710000
AP008214_137 13700001 13810000
AP008214_138 13800001 13910000
AP008214_139 13900001 14010000
AP008214_140 14000001 14110000
AP008214_141 14100001 14210000
AP008214_142 14200001 14310000
AP008214_143 14300001 14410000
AP008214_144 14400001 14510000
AP008214_145 14500001 14610000
AP008214_146 14600001 14710000
AP008214_147 14700001 14810000
AP008214_148 14800001 14910000
AP008214_149 14900001 15010000
AP008214_150 15000001 15110000
AP008214_151 15100001 15210000
AP008214_152 15200001 15310000
AP008214_153 15300001 15410000
AP008214_154 15400001 15510000
AP008214_155 15500001 15610000
AP008214_156 15600001 15710000
AP008214_157 15700001 15810000
AP008214_158 15800001 15910000
AP008214_159 15900001 16010000
AP008214_160 16000001 16110000
AP008214_161 16100001 16210000
AP008214_162 16200001 16310000
AP008214_163 16300001 16410000
AP008214_164 16400001 16510000
AP008214_165 16500001 16610000
AP008214_166 16600001 16710000
AP008214_167 16700001 16810000
AP008214_168 16800001 16910000
AP008214_169 16900001 17010000
AP008214_170 17000001 17110000
AP008214_171 17100001 17210000
AP008214_172 17200001 17310000
AP008214_173 17300001 17410000
AP008214_174 17400001 17510000
AP008214_175 17500001 17610000
AP008214_176 17600001 17710000
AP008214_177 17700001 17810000
AP008214_178 17800001 17910000
AP008214_179 17900001 18010000
AP008214_180 18000001 18110000
AP008214_181 18100001 18210000
AP008214_182 18200001 18310000
AP008214_183 18300001 18410000
AP008214_184 18400001 18510000
AP008214_185 18500001 18610000
AP008214_186 18600001 18710000
AP008214_187 18700001 18810000
AP008214_188 18800001 18910000
AP008214_189 18900001 19010000
AP008214_190 19000001 19110000

Query Match      80.0%; Score 16; DB 15; Length 110000;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      5  TGGTTGGGCAACACAT 20
      |||||
Db      67577 TGGTTGGGCAACACAT 67592

RESULT 44
AC097534
LOCUS      AC097534      139455 bp      DNA      linear      PRI 01-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-798M19 from 4, complete sequence.
ACCESSION      AC097534      AC022744
VERSION      AC097534.3      GI:16874920
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 139455)
AUTHORS      Sulston,J.E. and Waterston,R.
TITLE      Toward a complete human genome sequence
JOURNAL      Genome Res. 8 (11), 1097-1108 (1998)
PUBMED      9847074
REFERENCE      2 (bases 1 to 139455)
AUTHORS      Cedroni,M., Haakenson,W. and Creason,K.
TITLE      The sequence of Homo sapiens BAC clone RP11-798M19
JOURNAL      Unpublished (2001)
REFERENCE      3 (bases 1 to 139455)
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (18-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE      4 (bases 1 to 139455)
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (09-NOV-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE      5 (bases 1 to 139455)
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE      6 (bases 1 to 139455)
AUTHORS      Waterston,R.
TITLE      Direct Submission
JOURNAL      Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 9, 2001 this sequence version replaced gi:16604096.
COMMENT      ----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0798M19
Drafting Center: WIBR
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
```

clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

# MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

# SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,B., Tateno,M., Catanesi,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBAC3.6

# NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-500A5; the clone sequenced to the right is RP11-489G11, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-798M19; actual end is at base position 37652 of RP11-489G11.

Data from AC09588 and AC022691 was used to finish this clone, AC097534.

The sequence of AC022744 has been incorporated into AC097534.

## FEATURES

Location/Qualifiers

1..139455  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/map="4"  
/clone="RP11-798M19"  
/clone\_lib="RPCI-11"

2510..2548  
/rpt\_family="TCCC)n"  
2567..2608  
/rpt\_family="(TTTC)n"  
3513..3543  
/rpt\_family="AT\_rich"  
4559..4739  
/rpt\_family="MIR"  
4897..5120  
/rpt\_family="MER1\_type"  
5609..5913  
/rpt\_family="Alu"  
6139..6216  
/rpt\_family="MER2\_type"  
6217..6689  
/rpt\_family="ERVK"  
6690..7435  
/rpt\_family="MER2\_type"  
6811..6832  
/rpt\_family="AT\_rich"  
7512..7819  
/rpt\_family="Alu"  
7796..7824  
/rpt\_family="AT\_rich"  
7844..9184  
/rpt\_family="MER1\_type"  
7886..7980

repeat\_region  
10470..10509  
/rpt\_family="MIR"  
11543..11602  
/rpt\_family="ERVK"  
11565..11610  
/rpt\_family="(TC)n"  
11622..11691  
/rpt\_family="CT-rich"  
11935..11952  
/rpt\_family="MIR"  
11953..12364  
/rpt\_family="Alu"  
12265..12409  
/rpt\_family="MIR"  
12671..12856  
/rpt\_family="L1"  
12846..12904  
/rpt\_family="(TG)n"  
12896..13198  
/rpt\_family="L1"  
13717..13815  
/rpt\_family="MIR"  
14135..14212  
/rpt\_family="Mariner"  
14213..14498  
/rpt\_family="MER2\_type"  
14508..14534  
/rpt\_family="(T)n"  
14673..14689  
/note="match to EST AI207460 (NID:g6361468)"  
14675..14689  
/note="match to EST BG505352 (NID:g13466856)"  
17415..17448  
/rpt\_family="(TTTA)n"  
17419..17732  
/rpt\_family="Alu"  
17787..17855  
/rpt\_family="AT\_rich"  
17828..18065  
/rpt\_family="L1"  
18933..19073  
/rpt\_family="CR1"  
20455..20660  
/rpt\_family="MIR"  
20561..20590  
/rpt\_family="AT\_rich"  
23002..23033  
/rpt\_family="(ATTG)n"  
23072..23106  
/rpt\_family="AT\_rich"  
24108..24794  
/rpt\_family="CR1"  
25049..25457  
/rpt\_family="MaLR"  
25644..25664  
/rpt\_family="AT\_rich"  
25668..25970  
/rpt\_family="Alu"  
27131..27319  
/note="match to EST BE837547 (NID:g10269925)"  
27855..28073  
/rpt\_family="ERV1"  
28125..28175  
/rpt\_family="ERV1"  
28153..28251  
/rpt\_family="ERV1"  
28256..28366  
/rpt\_family="ERV1"  
29347..29634  
/rpt\_family="Alu"  
33123..33294

Query Match	80.0%;	Score 16;	DB 8;	Length 139455;
Best Local Similarity	100.0%;	Pred. No. 44;		
Matches	16;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
			0;	
QY	4	TTGGTTGGGCAACACA	19	
Db	121340	TTGGTTGGGCAACACA	121355	
RESULT 45				
AP005741				
LOCUS				
DEFINITION	AP005741	141257 bp	DNA	linear
ACCESSION	Oryza sativa (japonica cultivar-group)	genomic DNA, chromosome 8,		
VERSION	PAC clone:P0470H09.			
KEYWORDS	AP005741.3	GI:37573143		
SOURCE	Oryza sativa (japonica cultivar-group)			
ORGANISM	Oryza sativa (japonica cultivar-group)			
	Eukaryota; Viridiplantae; Streptophyta;			
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
	Ehrhartoideae; Oryzeae; Oryza.			
REFERENCE	1			
AUTHORS	Sasaki,T., Matsumoto,T. and Katayose,Y.			
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC			
	clone:P0470H09			
JOURNAL	Published Only in Database (2002)			
REFERENCE	2	(bases 1 to 141257)		
AUTHORS	Sasaki,T., Matsumoto,T. and Katayose,Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (11-SEP-2002) Takuji Sasaki, National Institute of			
	Agrobiological Sciences, Rice Genome Research Program; Kannondai			
	2-1-2, Tsukuba, Ibaraki 305-8602, Japan			
	(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,			
	Tel.81-298-38-7411, Fax:81-298-38-7468)			
COMMENT	On Oct 7, 2003 this sequence version replaced gi:30984152.			
	Genes were predicted from the integrated results of the following:			
	GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH			
	(http://www.softberry.com/), GeneMark.hmm			
	(http://opal.biology.gatech.edu/GeneMark/), GlimmerM			
	(http://www.tigr.org/tdb/glimmer/glmr_form.html), RiceHMM			
	(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor			
	(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4			
	(http://globin.cse.psu.edu/html/docs/sim4.html), gap2			
	(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The			
	genomic sequence was searched against NCBI NonRedundant Protein			
	database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA			
	sequence database at RGP or DBJ. Protein homologues of the coding			
	regions were searched against NCBI NonRedundant Protein database			
	with BLASTP. ESTs represent the identified cDNA sequences using			
	BLASTN with the corresponding DBJ accession no. and RGP clone ID.			
	Full-length cDNAs represent the identified cDNA sequences using			
	BLASTN with the corresponding DBJ accession no.			
	A gene with identity or significant homology to a protein is			
	classified based on the protein name to indicate the homology level			
	such as same name, 'putative-' and '-like protein'. A gene without			
	significant homology to any protein but with full-length cDNA or			
	EST homology (covering almost the entire length of partial			
	sequence) is classified as an 'unknown' protein. A gene predicted			
	by two or more gene prediction programs is classified as a			
	'hypothetical' protein according to IRGSP standard. A gene			
	predicted by a single gene prediction program is also classified as			
	a probable 'hypothetical' protein and is included as a			
	miscellaneous feature of the sequence.			
	The orientation of the sequence is from SP6 to T7 of the PAC clone.			
	This sequence of P0470H09 clone has an overlap with P0571B09 (DBJ:			
	AP005526) clone at 5' end and with P0015A04 (DBJ: AP005740) clone			
	at 3' end. Detailed information on overlap and assembly quality			
	together with annotation of this entry is available at			
	http://rgp.dna.affrc.go.jp/GenomeSeq.html.			
FEATURES	Location/Qualifiers			
source	1..141257			
	/organism="Oryza sativa (japonica cultivar-group)"			



```

misc_feature
21602..24700
/gene="P0470H09.7"
/note="GAG-POL precursor RIRE2"
gene
complement(25017..25478)
/gene="P0470H09.8"
misc_feature
complement(25017..25478)
/gene="P0470H09.8"
/note="putative orf5 RIRE2"
gene
complement(26210..26977)
/gene="P0470H09.9"
misc_feature
complement(26210..26977)
/gene="P0470H09.9"
/note="putative orf4 RIRE2"
gene
complement(27064..27960)
/gene="P0470H09.10"
misc_feature
complement(27064..27960)
/gene="P0470H09.10"
/note="putative orf3 RIRE2"
gene
28982..29422
/gene="P0470H09.6"
LTR
28982..29422
/gene="P0470H09.6"
/note="3' LTR"
gene
30371..34991
/gene="P0470H09.11"
mRNA
join(<30371..30914,31173..31337,32008..32093,32908..33047,
33670..34279,34905..34991)
/gene="P0470H09.11"
/note="start and end point are not identified"
CDS
join(30371..30914,31173..31337,32008..32093,32908..33047,
33670..34279,34905..34991)
/gene="P0470H09.11"
/codon_start=1
/product="circumsporozoite protein-like protein"
/protein_id="BAD33858.1"
/db_xref="GI:50726282"
translation="MALAARAGELHNLGPAADSFCAARSPSWSTARVGAASP
SRPPSGHTSPAAARAAAEAEELGEVRKLNRRRAVSAYRRKMPARGAGQRPV
DDLDVAERHLLGVLPIRLGQVPALVPRYLAVASSKNTARMSNIERRIGL
SPWEAAKSNCKEYKSTRWRVFTCTDVLSEHNDMPPEQPMKTHASMLAHSR
RKGRPPASAGAKDRAGLLRELLVLSGIRFIPFNSWGGYLRARFASDPDVGS
KTFVKPYPSFSDPEKEDRRYPCRVHDYGRASRPGPSRVVPGTGLSATRLTVSR
VHAGARVGGRCGACFAEDAVRAHAQPHSSRWTCACAREGKRELRASADPVAA
EVATPRLRGSHGRVEEDAGNRKRTAVASSGANGDTSSEGHETGLMKTGDEP
TARINQMLDGGILRRQPAAREGGNGDGRFGRARASTRLRESVAGVGLGGATP
SEAGDERASAGCGAEAMTTAMTAGRFGAERRHQ"
complement(38381..42196)
/gene="P0470H09.12"
misc_feature
complement(38381..42196)
/note="probably inactive due to including stop codon(s) in
CDS
pseudogene, Fourf gag/pol protein"
/pseudo
53728..54078
/gene="P0470H09.13"
mRNA
<53728..54078
/gene="P0470H09.13"
/note="start and end point are not identified"
CDS
53728..54078
/gene="P0470H09.13"
/note="predicted by GlimmerM etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD33859.1"

Query Match      80.0%; Score 16; DB 15; Length 141257;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      5 TGGTTGGGCAACACAT 20
|||||
Db      45858 TGGTTGGGCAACACAT 45873

```

```

RESULT 46
AC093739/c
LOCUS
DEFINITION      Homo sapiens chromosome 2 clone RP11-22A4, WORKING DRAFT SEQUENCE,
2 unordered pieces.
AC093739 AC019029
VERSION
AC093739.1 GI:15529777
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 154612)
Waterston,R.H.
Direct Submission
Submitted (10-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 10, 2001 this sequence version replaced gi:7230152.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
Center project name: H_NH0022A04
----- Project Information -----
Drafting center: WIBR
----- Summary Statistics -----
Sequencing vector: M13; 52k
Sequencing vector: plasmid; 47k
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 99% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 152111 bases at least Q40
Consensus quality: 152371 bases at least Q30
Consensus quality: 154464 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
*
* 1 54124: contig of 54124 bp in length
* 54125 54224: gap of unknown length
* 54225 154612: contig of 100388 bp in length.
* Location/Qualifiers
* 1. 154612
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="2"
* /clone="RP11-22A4"
* 1. 54124
* /note="assembly_name:Contig8
* clone_end:17
* vector_side:right"
* 54125..54224
* /estimated_length=unknown
* 54225..154612
* /note="assembly_name:Contig9
* clone_end:SP6

FEATURES
source
1. 154612
misc_feature
1. 54124
/note="assembly_name:Contig8
clone_end:17
vector_side:right"
gap
54125..54224
/estimated_length=unknown
misc_feature
54225..154612
/note="assembly_name:Contig9
clone_end:SP6

```

## ORIGIN vector\_side:left"

Query Match 80.0%; Score 16; DB 14; Length 154612;  
Best Local Similarity 100.0%; Pred.No. 44;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCTTTGGTTGGGCAAC 16  
|||||

Db 113017 GCTTTGGTTGGGCAAC 113002  
|||||

RESULT 47  
AC009311/c  
LOCUS 160625 bp DNA linear PRI 21-APR-2005  
DEFINITION Homo sapiens BAC clone RP11-311L3 from 2, complete sequence.

AC009311  
AC009311  
VERSION AC009311.3 GI:10716649

HTG.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 160625)

Nguyen,C., Maupin,R., Hawkins,M. and Le,T.

The sequence of Homo sapiens BAC clone RP11-311L3

Unpublished (2001)

2 (bases 1 to 160625)

Waterston,R.H.

Direct Submission

Submitted (13-AUG-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

3 (bases 1 to 160625)

Waterston,R.H.

Direct Submission

Submitted (15-MAR-2000) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

4 (bases 1 to 160625)

Waterston,R.

Direct Submission

Submitted (07-OCT-2000) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

5 (bases 1 to 160625)

Waterston,R.H.

Direct Submission

Submitted (25-MAR-2001) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

6 (bases 1 to 160625)

Waterston,R.

Direct Submission

Submitted (29-OCT-2002) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

7 (bases 1 to 160625)

Wilson,R.K.

Direct Submission

Submitted (21-APR-2005) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Oct 7, 2000 this sequence version replaced gi:7243908.

COMMENT

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)

----- Summary Statistics

-----

Center project name: H\_NH0311L03

-----

NOTICE:

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,  
Department of Genetics, Washington University, St. Louis MO. For  
additional information about the map position of this sequence, see  
<http://genome.wustl.edu>

## SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male  
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,  
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved  
approach for construction of bacterial artificial chromsome  
libraries, Genomics 51:1-8. The clone may be obtained either from  
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong  
and coworkers at <http://www.chori.org>  
VECTOR: pBACe3.6

## NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-52K24, 200 base pair  
overlap. The clone sequenced to the right is RP11-502N7, 200 base  
pair overlap. Actual start of this clone is at base position 1 of  
RP11-311L3; actual end is at base position 160625 of RP11-311L3.

## FEATURES

	Location/Qualifiers
source	1..160625 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="2" /clone="RP11-311L3" /clone_lib="RPCI-11" 178..773 /note="CpG island (dGC=67.8, o/e=0.80, #CpGs=56)" complement(9279..28668) /gene="NMI"
misc_feature	complement(join(9279..9683,10434..10540,14292..14478, 14566..14672,17636..17798,20756..20851,21676..21762, 28401..28668)) /gene="NMI"
gene	complement(join(9501..9683,10434..10540,14292..14478, 14566..14672,17636..17798,20756..20851,21676..21756)) /gene="NMI"
CDS	mrna: H_NH0311L03.1 This gene was based on gi(4758813)" /codon_start=1 /product="unknown" /protein_id="AAV15066.1" /db_xref="GI:62822518" translation="MEADKDDTOILKEHSPDEFIKDQNKGLIDITTKNIOLKKEI QKLELQEAKEFQIKEDIPETKMKPLSVETPENSQLSNCSFQVSKVPYELQK QALTEPKEEVAAQNVVSMKHVQIQNVLEVPVPLNSGVRFPVYVEVSKMKN VTSEIPLTREDQDKLELSFSGRNGGGVDVDRQSGSAVITFEIGVADKILK KREYPIYINQTRHRTVSPYTEIHLKKYQIFSGTSKRTVLLTGMGEGIQMDEEVEDLI NIHFQAKNGGGEVDVVKSLGQPHIAYFEE"
misc_feature	27754..28613 /note="CpG island (dGC=62.9, o/e=0.79, #CpGs=72)"
misc_feature	34781..35193 /note="CpG island (dGC=74.1, o/e=0.66, #CpGs=40)"
gene	96399..118854 /gene="TNFAIP6"
mrna	108828..109056,112357..112397,118172..118854 /gene="TNFAIP6"
CDS	108828..109056,112357..112397,118172..118341 108828..109056,112357..102888,104864..105025, join(96399..96568,102751..102888,104864..105025, 108828..109056,112357..112397,118172..118854), join(96475..96568,102751..102888,104864..105025, 108828..109056,112357..112397,118172..118341)

```
/gene="TNFAIP6"
/note="Homo sapiens tumor necrosis factor, alpha-induced
protein 6 (TNFAIP6), mRNA.; H NH0311L03.2
This gene was based on gi(26051242)"
/codon_start=1
/product="unknown"
/protein_id="AAV15067.1"
/db_xref="GI:62822519"
/translation="MILLIYLLWEDTQGFQKDGIFHNSIWLRAAGVYHRRARS
GKYUTYAEKAVCFEGGHLATYKLEARKIGFHVCAAGMAKRGVYIVKPGPN
CFGKGTIGIDYILNRERWDAYCYNPAKCEGGVFDPKQIPKSPGPFNVEYDQNI
CYWHLRYQGRHLSFLDFDLEDDPGCLADYVEIYDSYDVHGVGRYCGDELDDI
ISTGNVMTLKFLSDASVTAGGQFIKYVAMDPSKSSQKQNTSTTGNKNFLAGRFH
L"
misc_feature 147349..147620
/note="CpG island (%GC=61.4, o/e=0.83, #CpGs=20)"
148566..149580
/note="CpG island (%GC=65.1, o/e=0.88, #CpGs=84)"
ORIGIN
Query Match 80.0%; Score 16; DB 8; Length 160625;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAAC 16
|||||
Db 121737 GCTTTGGTTGGGCAAC 121722

RESULT 48
AC009588/c
LOCUS AC009588 166472 bp DNA linear HTG 09-SEP-2000
DEFINITION Homo sapiens chromosome 4 clone RP11-297M24 map 4, WORKING DRAFT
SEQUENCE, 18 unordered pieces.
ACCESSION AC009588
VERSION 5 GI:10045234
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 166472)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collamore,A.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McSwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo.A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (27-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2000 this sequence version replaced gi:6850482.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIER
Web site: http://www-seq.wi.mit.edu

FEATURES
source
1..166472
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-297M24"
/clone_lib="RPC1-11 Human Male BAC"
1..17576
/note="assembly_fragment
clone_end:SP6
vector_side:left"
17577..17676
gap
```

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1737

Center clone name: 297\_M\_24

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator-amerham; 4% of reads

Chemistry: Dye-terminator Big Dye; 96% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156962 bases at least Q40

Consensus quality: 161562 bases at least Q30

Consensus quality: 163407 bases at least Q20

Insert size: 173000; agarose-fp

Insert size: 164772; sum-of-contigs

Quality coverage: 4.5 in Q20 bases; agarose-fp

Quality coverage: 4.7 in Q20 bas.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 18 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 17576: contig of 17576 bp in length

\* 17577 17676: gap of 100 bp

\* 17677 19324: contig of 1648 bp in length

\* 19325 19424: gap of 100 bp

\* 19425 59066: contig of 39642 bp in length

\* 59067 59166: gap of 100 bp

\* 59167 62314: contig of 3148 bp in length

\* 62315 62414: gap of 100 bp

\* 62415 65191: contig of 2777 bp in length

\* 65192 65291: gap of 100 bp

\* 65292 69715: contig of 4424 bp in length

\* 69716 69815: gap of 100 bp

\* 69816 74395: contig of 4580 bp in length

\* 74396 74495: gap of 100 bp

\* 74496 79903: contig of 5408 bp in length

\* 79904 80003: gap of 100 bp

\* 80004 85577: contig of 5574 bp in length

\* 85578 85677: gap of 100 bp

\* 85678 94570: contig of 8893 bp in length

\* 94571 94671: gap of 100 bp

\* 94671 101824: contig of 7154 bp in length

\* 101825 101924: gap of 100 bp

\* 101925 108382: contig of 6458 bp in length

\* 108383 108482: gap of 100 bp

\* 108483 117723: contig of 9241 bp in length

\* 117724 117823: gap of 100 bp

\* 117824 127631: contig of 9808 bp in length

\* 127632 127731: gap of 100 bp

\* 127732 138685: contig of 10954 bp in length

\* 138686 138785: gap of 100 bp

\* 138786 150114: contig of 11329 bp in length

\* 150115 150214: gap of 100 bp

\* 150215 164874: contig of 14660 bp in length

\* 164875 164974: gap of 100 bp

\* 164975 166472: contig of 1498 bp in length.

```
misc_feature /estimated_length=100
17677..19324
/note="assembly_fragment"
gap 19325..19424
/estimated_length=100
misc_feature 19425..59066
/note="assembly_fragment"
gap 59067..59166
/estimated_length=100
misc_feature 59167..62314
/note="assembly_fragment"
gap 62315..62414
/estimated_length=100
misc_feature 62415..65191
/note="assembly_fragment"
gap 65192..65291
/estimated_length=100
misc_feature 65292..69715
/note="assembly_fragment"
gap 69716..69815
/estimated_length=100
misc_feature 69816..74395
/note="assembly_fragment"
gap 74396..74495
/estimated_length=100
misc_feature 74496..79503
/note="assembly_fragment"
gap 79904..80003
/estimated_length=100
misc_feature 80004..85577
/note="assembly_fragment"
gap 85578..85677
/estimated_length=100
misc_feature 85678..94570
/note="assembly_fragment"
gap 94571..94670
/estimated_length=100
misc_feature 94671..101824
/note="assembly_fragment"
gap 101825..101924
/estimated_length=100
misc_feature 101925..108382
/note="assembly_fragment"
gap 108383..108482
/estimated_length=100
misc_feature 108483..117723
/note="assembly_fragment"
gap 117724..117823
/estimated_length=100
misc_feature 117824..127631
/note="assembly_fragment"
gap 127632..127731
/estimated_length=100
misc_feature 127732..138695
/note="assembly_fragment"
gap 138696..138785
/estimated_length=100
misc_feature 138786..150114
/note="assembly_fragment"
gap 150115..150214
/estimated_length=100
misc_feature 150215..164874
/note="assembly_fragment"
gap 164875..164974
/estimated_length=100
misc_feature 164975..166472
/note="assembly_fragment
clone end:T7
vector_side:right"
```

ORIGIN

Query Match 80.0%; Score 16; DB 14; Length 166472;  
Best Local Similarity 100.0%; Pred. No. 43;

```
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 TTGGTTGGCAACACA 19
|||||
Db 106452 TTGGTTGGCAACACA 106437

RESULT 49
BX537258 188776 bp DNA linear ROD 17-DEC-2004
LOCUS Mouse DNA sequence from clone RP23-152P13 on chromosome 4 Contains
DEFINITION a eukaryotic translation initiation factor 4E (Eif4e) pseudogene
and a eukaryotic translation initiation factor 4A (Eif4a)
pseudogene, complete sequence.
ACCESSION BX537258
VERSION BX537258.3 GI:45581000
KEYWORDS HTG; Eif4a; Eif4e.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 188776)
AUTHORS North,P., Leaves,N., Greystrong,J., Coppola,M., Manjunath,S.,
Russell,E., Smith,M., Strachan,G., Tofts,C., Boal,E., Cobley,V.,
Hunter,G., Kimberley,C., Thomas,D., Cave-Berry,L., Weston,P. and
Botcherby,M.R.M.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-2004) Mouse Sequencing Group, HGMP-RC, Hinxton,
Cambridge, CB10 1SB, UK. E-mail enquiries:- mzbotche@hgmp.mrc.ac.uk
or pnorth@hgmp.mrc.ac.uk
REMARK HGMP-RC part of the UK Mouse Sequencing Consortium
COMMENT On Mar 18, 2004 this sequence version replaced gi:32451244.
----- Genome Center
Center: UK Medical Research Council
Center code: UK-MRC
Web site: http://mrcseq.har.mrc.ac.uk
Contact: mouseq@har.mrc.ac.uk
-----
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/RP23-152P13
from the RPCI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
FEATURES
source
1..188776
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-152P13"
/clone_lib="RPCI-23"
23439..24091
/locus_tag="RP23-152P13.1-001"
/pseudo
23439..24091
/locus_tag="RP23-152P13.1-001"
/note="match: proteins: Sw:P06730 Sw:P20415 Sw:P29338"
gene
CDS
```

```

Sw:P48597 Sw:Q9N0T5 Tr:Q8C470 Tr:Q9DF56 Tr:Q9W7B5"
/pseudo
/codon_start=1
/product="eukaryotic translation initiation factor 4E
(Bif4e) pseudogene"
/complement(47863..49080)
/locus_tag="RP23-152P13.2-001"
/pseudo
/complement(47863..49080)
/locus_tag="RP23-152P13.2-001"
/Note="match: proteins: Sw:P04765 Sw:P29562 Tr:AAH49915
Tr:AAH63812 Tr:Q42576 Tr:Q7ZU67 Tr:Q7ZX50 Tr:Q802C9"
/pseudo
/codon_start=1
/product="eukaryotic translation initiation factor 4A
(Bif4a) pseudogene"

ORIGIN
Query Match      80.0%; Score 16; DB 9; Length 188776;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAAC 16
|||||
Db 131236 GCTTTGGTTGGGCAAC 131251

RESULT 50
AC116953
LOCUS
DEFINITION AC116953 193506 bp DNA linear ROD 25-NOV-2003
Mus musculus BAC clone RP23-119H17 from chromosome 7, complete
sequence.
ACCESSION AC116953
VERSION AC116953.5 GI:37202265
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 193506)
AUTHORS VanBrunt,A.
TITLE The sequence of Mus musculus BAC clone RP23-119H17
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 193506)
AUTHORS Wilson,R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 193506)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (03-APR-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 193506)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 193506)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 193506)
AUTHORS Wilson,R.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 1, 2003 this sequence version replaced gi:34495087.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu

```

Contact: submissions@watson.wustl.edu  
----- Summary Statistics  
Center project name: M\_BA0119H17  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

#### SOURCE INFORMATION:

The RPCI-23 BAC library has been constructed by Kazutoyo Osegawa and Minako Tateo in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

#### NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES	source
repeat_region	1. 193506
repeat_region	/organism="Mus musculus"
repeat_region	/mol_type="genomic DNA"
repeat_region	/db_xref="taxon:10090"
repeat_region	/chromosome="7"
repeat_region	/map="7"
repeat_region	/clone="RP23-119H17"
repeat_region	/clone_lib="RPCI-23"
repeat_region	1188..1337
repeat_region	/rpt_family="MaLR"
repeat_region	1751..1854
repeat_region	/rpt_family="ERV1"
repeat_region	2287..2340
repeat_region	/rpt_family="MaLR"
repeat_region	3713..4060
repeat_region	/rpt_family="L1"
repeat_region	4250..4263
repeat_region	/note="Sequence derived from one plasmid subclone."
repeat_region	4356..4482
repeat_region	/rpt_family="ERVK"
repeat_region	4768..4951
repeat_region	/rpt_family="ERVK"
repeat_region	6342..6676
repeat_region	/rpt_family="ERV1"
repeat_region	7528..7833
repeat_region	/rpt_family="MaLR"
repeat_region	7966..8021
repeat_region	/rpt_family="ERV1"
repeat_region	9226..9318
repeat_region	/rpt_family="Alu"
repeat_region	9488..9756
repeat_region	/rpt_family="RMR15"
repeat_region	14323..14847
repeat_region	/rpt_family="L1"
repeat_region	14951..15141
repeat_region	/rpt_family="L1"
repeat_region	15341..15669
repeat_region	/rpt_family="L1"

```
repeat_region 16000. .16253 /rpt family="MER2_type"
repeat_region 16390. .17076 /rpt family="L1"
repeat_region 17070. .17582 /rpt family="L1"
repeat_region 17583. .17639 /rpt family="L1"
repeat_region 18980. .19109 /rpt family="B4"
repeat_region 20854. .21257 /rpt family="MaLR"
repeat_region 21308. .21513 /rpt family="B4"
repeat_region 21950. .22267 /rpt family="MaLR"
repeat_region 23809. .24149 /rpt family="MaLR"
repeat_region 25091. .25401 /rpt family="L1"
repeat_region 25648. .25711 /rpt family="ERV1"
repeat_region 25788. .26206 /rpt family="L1"
repeat_region 27178. .27986 /rpt family="L1"
repeat_region 28187. .28342 /rpt family="Alu"
repeat_region 28818. .29334 /rpt family="RMR15"
repeat_region 29529. .29849 /rpt family="MaLR"
repeat_region 30823. .31686 /rpt family="L1"
repeat_region 31886. .32303 /rpt family="L1"
repeat_region 32462. .32655 /rpt family="B2"
repeat_region 33109. .33317 /rpt family="ERVK"
repeat_region 33900. .34109 /rpt family="ERVK"
repeat_region 34446. .34629 /rpt family="ERV1"
repeat_region 34719. .35661 /rpt family="L1"
repeat_region 35708. .35872 /rpt family="ERV1"
repeat_region 38220. .38492 /rpt family="B4"
repeat_region 38880. .39273 /rpt family="MaLR"
repeat_region 39577. .39893 /rpt family="ERVK"
repeat_region 41222. .41609 /rpt family="L1"
repeat_region 41995. .42328 /rpt family="MaLR"
repeat_region 42336. .42880 /rpt family="MaLR"
repeat_region 42913. .43044 /rpt family="ERV1"
repeat_region 43052. .43119 /rpt family="MaLR"
repeat_region 43120. .43511 /rpt family="MaLR"
repeat_region 43512. .43614 /rpt family="MaLR"
repeat_region 43627. .43884 /rpt family="L1"
repeat_region 43905. .44367 /rpt family="MaLR"
repeat_region 44823. .45209
```

```
repeat_region /rpt family="MaLR"
repeat_region /rpt family="Alu"
repeat_region /rpt family="ERV1"
repeat_region /rpt family="RMR1A"
repeat_region /rpt family="ERV1"
repeat_region /rpt family="ERV1"
```

Query Match 80.0%; Score 16; DB 9; Length 193506;  
Best Local Similarity 100.0%; Pred.No. 43;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTGGTTGGCAACA 17  
Db 138492 CTTGGTTGGCAACA 138507  
|||||

Search completed: May 8, 2006, 04:49:00  
Job time : 1875 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2006, 00:53:47 ; Search time 675 Seconds  
(without alignments)  
197.472 Million cell updates/sec

Title: US-10-671-074-172

Perfect score: 20

Sequence: 1 gcttggttggaacacacat 20

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size : 1

Total number of hits satisfying chosen parameters: 9993364

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N Geneseq 21.\*

1: Geneseqm1980s.\*

2: Geneseqm1990s.\*

3: Geneseqm2000s.\*

4: Geneseqm2001as.\*

5: Geneseqm2001bs.\*

6: Geneseqm2002as.\*

7: Geneseqm2002bs.\*

8: Geneseqm2003as.\*

9: Geneseqm2003bs.\*

10: Geneseqm2003cs.\*

11: Geneseqm2003ds.\*

12: Geneseqm2004as.\*

13: Geneseqm2004bs.\*

14: Geneseqm2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	20	100.0	20 12	ADN97882
2	20	100.0	20 12	ADN40818
3	20	100.0	254 5	AS14337
4	20	100.0	520 10	ADK61538
5	20	100.0	578 10	ADP79645
6	20	100.0	2514 5	AS85052
7	20	100.0	2827 10	ACCT72829
8	20	100.0	3200 12	ADQ03121
9	20	100.0	3200 12	ADQ03123
10	20	100.0	3421 13	ADR83528
11	20	100.0	3517 6	ABS73280
12	20	100.0	3517 12	ADQ03122
13	20	100.0	4945 12	ADN97721
14	20	100.0	4945 12	ADN40657
15	20	100.0	5722 13	ACF87489
16	20	100.0	5722 13	ACF87572
17	20	100.0	5723 6	ABV94143
18	20	100.0	5723 12	ADN06031
19	20	100.0	5723 12	ADN97714

c 20	20	100.0	5723	12	ADN40650	Adn40650 Human for
c 21	20	100.0	5723	13	ACN39628	Acn39628 Tumour-as
c 22	20	100.0	5723	14	ADY14881	Ady14881 DNA encod
c 23	20	100.0	5769	10	ADF81591	Adf81591 Leukaemia
c 24	20	100.0	5833	5	AS85053	As85053 DNA encod
c 25	19	95.0	20	12	ADN97750	Adn97750 Human fox
c 26	19	95.0	20	12	ADN97828	Adn97828 Human fox
c 27	19	95.0	20	12	ADN40686	Adn40686 Human fox
c 28	19	95.0	20	12	ADN40764	Adn40764 Human for
c 29	17	85.0	2000	6	AB216434	Ab216434 Arabidops
c 30	16	80.0	20	12	ADN97751	Adn97751 Human fox
c 31	16	80.0	20	12	ADN97829	Adn97829 Human fox
c 32	16	80.0	20	12	ADN40687	Adn40687 Human for
c 33	16	80.0	20	12	ADN40765	Adn40765 Human for
c 34	16	80.0	4106	4	ABL05051	Aal05051 Human rep
c 35	16	80.0	4106	4	ABL97944	Ab197944 Human tes
c 36	16	80.0	160198	10	ADL13962	Adl13962 Osteoarth
c 37	16	80.0	164772	10	ADL13904	Adl13904 Osteoarth
c 38	15	75.0	406	12	ADL88066	Adl88066 DNA up-re
c 39	15	75.0	406	12	ADL88067	Ach45273 Human foe
c 40	15	75.0	498	9	ACH45273	Ach45273 Human foe
c 41	15	75.0	574	4	AAK88397	Aak88397 Human dig
c 42	15	75.0	574	5	AS339498	As339498 cDNA enco
c 43	15	75.0	574	9	ADB32224	Adb32224 Human nov
c 44	15	75.0	1263	5	AS75868	As75868 DNA encod
c 45	15	75.0	1342	2	AAV31252	Aav31252 E. coli J
c 46	15	75.0	1550	11	ADM02861	Adm02861 Human cDN
c 47	15	75.0	1635	6	ABQ75898	Abq75898 Human ubi
c 48	15	75.0	1701	10	ADF03345	Adf03345 Bacterial
c 49	15	75.0	2154	8	ACA21214	Ac21214 Prokaryot
c 50	15	75.0	2181	9	ADA32702	Ada32702 DNA encod
c 51	15	75.0	2550	4	AAH14878	Aah14878 Human cDN
c 52	15	75.0	3059	10	ADA53348	Ada53348 Human cod
c 53	15	75.0	3342	4	ABL23542	Ab123542 Drosophill
c 54	15	75.0	3366	4	ABL17225	Ab117225 Drosophill
c 55	15	75.0	3593	4	AAD02709	Aad02709 Murine B7
c 56	15	75.0	3593	5	AAD02774	Aad02774 Murine B7
c 57	15	75.0	3593	10	AAD51099	Aad51099 Murine B7
c 58	15	75.0	3593	10	ABT14004	Abt14004 Murine PD
c 59	15	75.0	3593	12	ADJ75888	Adj75888 Marker ge
c 60	15	75.0	3593	14	ADZ61822	Adz61822 Murine Pd
c 61	15	75.0	5205	12	ADJ95458	Adj95458 Human cDN
c 62	15	75.0	5205	13	ADS86880	Ad86880 cDNA enco
c 63	15	75.0	5205	14	ADZ75838	Adz75838 Human E3a
c 64	15	75.0	5466	11	AEA90301	Aea90301 Human NOV
c 65	15	75.0	5492	4	ABL17224	Ab117224 Drosophill
c 66	15	75.0	6308	12	ADJ95441	Adj95441 Human cDN
c 67	15	75.0	6308	13	ADS86863	Ad86863 cDNA enco
c 68	15	75.0	6308	14	ADZ75821	Adz75821 Human E3a
c 69	15	75.0	7648	12	AD116377	Ad116377 Human pro
c 70	15	75.0	7742	8	ABZ24689	Abz24689 Human cel
c 71	15	75.0	110000	2	AAT42063	Continuation (12 o
c 72	15	75.0	349980	13	ADT05649	Adt05649 Haemophil
c 73	14	70.0	20	12	ADP76896	Adp76896 Chimeric
c 74	14	70.0	20	12	ADP76326	Adp76326 Chimeric
c 75	14	70.0	20	12	ADP76474	Adp76474 Chimeric
c 76	14	70.0	20	12	ADP76296	Adp76296 Chimeric
c 77	14	70.0	20	12	ADP76342	Adp76342 Chimeric
c 78	14	70.0	20	12	ADP76371	Adp76371 Chimeric
c 79	14	70.0	20	12	ADP76928	Adp76928 Chimeric
c 80	14	70.0	88	10	ADB75728	Adb75728 Tomato pl
c 81	14	70.0	229	12	ADQ04357	Adq04357 Maize tra
c 82	14	70.0	231	12	ADQ04355	Adq04355 Maize tra
c 83	14	70.0	266	12	ADQ04374	Adq04374 Maize tra
c 84	14	70.0	296	12	ACH90349	Ach90349 Human gen
c 85	14	70.0	306	4	ABL21351	Ab121351 Drosophill
c 86	14	70.0	339	6	ABN16060	Abn16060 Human ORF
c 87	14	70.0	349	3	AAC22674	Aac22674 Human sec
c 88	14	70.0	404	12	ADQ04378	Adq04378 Maize tra
c 89	14	70.0	410	3	AAZ61682	Aaz61682 cDNA enco
c 90	14	70.0	410	4	AAZ99615	Aaz99615 Skin cell
c 91	14	70.0	419	12	ADQ04384	Adq04384 Maize tra
c 92	14	70.0	440	6	ABL34767	Ab134767 Murine cD

c 93	14	70.0	471	9	ACH43968	Ach43968 Human foe	166	14	70.0	2439	10	ADB72460	Adb72460 Human NFA
c 94	14	70.0	504	12	ACH76649	Ach76649 Human gen	167	14	70.0	2439	10	ADB95970	Ade95970 Human NFA
c 95	14	70.0	593	13	ACN60177	Acn60177 Cotton gy	c 168	14	70.0	2474	2	AAV65248	AAV65248 DNA encod
c 96	14	70.0	593	14	ACL55202	ACL55202 Human col	169	14	70.0	2478	14	ADV43378	Adv43378 Human psy
c 97	14	70.0	632	5	ABV57221	Abv57221 Human pro	170	14	70.0	2586	10	ADJ56367	Adj56367 Human cdn
c 98	14	70.0	657	4	AAH07975	Aah07975 Human cdn	171	14	70.0	2743	2	AAT36867	Aat36867 Human tra
c 99	14	70.0	723	10	ADB76212	Adb76212 Mouse Con	172	14	70.0	2743	4	AAH02897	Aah02897 Human ghe
100	14	70.0	748	13	ADQ82535	Adq82535 Wild type	173	14	70.0	2743	8	ACC83319	Acc83319 NF-ATc is
101	14	70.0	750	6	ABK24137	Abk24137 DNA encod	174	14	70.0	2749	2	AAQ86687	AAQ86687 Human NF-
102	14	70.0	750	6	ABK24136	Abk24136 DNA encod	175	14	70.0	2750	3	AAA23252	AAA23252 Human nuc
103	14	70.0	750	6	ABK24141	Abk24141 DNA encod	176	14	70.0	2751	5	AAF31710	AAF31710 Human NF-
104	14	70.0	750	6	ABK24127	Abk24127 DNA encod	177	14	70.0	2751	9	ADJ34048	Adj34048 Human cyt
105	14	70.0	750	6	ABK24130	Abk24130 DNA encod	178	14	70.0	2751	9	ADA66407	Ada66407 Human NF-
106	14	70.0	753	13	ADQ82527	Adq82527 Wild type	179	14	70.0	2793	8	ABZ79982	Abz79982 Human NF-
107	14	70.0	753	13	ADQ82525	Adq82525 Wild type	180	14	70.0	2853	2	AAQ84301	AAQ84301 Human NF-
108	14	70.0	753	13	ADQ82515	Adq82515 Wild type	181	14	70.0	2935	3	AAA35062	AAA35062 Human ade
109	14	70.0	753	13	ADQ82524	Adq82524 Wild type	182	14	70.0	2935	3	AAF21184	AAF21184 Human low
110	14	70.0	801	12	ADO63486	Ado63486 Transcrip	183	14	70.0	2935	9	ADA02721	Ada02721 Human NFA
111	14	70.0	816	13	ADS49135	Ads49135 Bacterial	184	14	70.0	2935	10	ADB72459	Adb72459 Human NFA
112	14	70.0	825	10	ADB76210	Adb76210 Mouse Con	185	14	70.0	2935	10	ADB95969	Ade95969 Human NFA
113	14	70.0	845	4	AAK35745	Aal35745 Human mus	186	14	70.0	2935	10	ABZ96878	Abz96878 Human nuc
114	14	70.0	887	4	AAK35747	Aal35747 Human mus	187	14	70.0	2935	11	ABD20727	Abd20727 Human pul
115	14	70.0	887	8	ABX58733	Abx58733 cDNA enco	188	14	70.0	3036	13	ADS48001	Ads48001 Bacterial
116	14	70.0	887	8	ABX58735	Abx58735 cDNA enco	189	14	70.0	3082	3	AAA30882	AAA30882 Human GFA
117	14	70.0	887	12	ADJ29485	Adj29485 Human mus	190	14	70.0	3082	8	ABA00846	AbA00846 GFAT rela
118	14	70.0	887	12	ADJ29483	Adj29483 Human mus	191	14	70.0	3082	8	ABA00845	AbA00845 GFAT rela
119	14	70.0	921	4	AAK35746	Aal35746 Human mus	192	14	70.0	3082	10	ADB31352	Adb31352 Bicalutam
120	14	70.0	921	8	ABX58734	Abx58734 cDNA enco	193	14	70.0	3082	10	ADE38372	Ade38372 Human pro
121	14	70.0	921	12	ADJ29484	Adj29484 Human mus	194	14	70.0	3082	12	ADP79271	Adp79271 Human GFA
122	14	70.0	1083	10	ADB76208	Adb76208 Mouse Con	195	14	70.0	3082	12	ADQ15105	Adq15105 Human can
c 123	14	70.0	1157	5	AA900006	Aas90006 DNA encod	196	14	70.0	3082	13	ADQ88169	Adq88169 Human 923
124	14	70.0	1194	13	ADR91659	Adr91659 Novel S.	197	14	70.0	3082	13	ADP23395	Adp23395 PRO polyp
125	14	70.0	1194	14	ADK45754	Adk45754 Streptoco	198	14	70.0	3082	14	ADZ70505	Adz70505 Human cdn
126	14	70.0	1194	14	AEA55529	Aea55529 Streptoco	199	14	70.0	3082	14	ADZ75427	Adz75427 Human glu
c 127	14	70.0	1275	8	ADA69899	Ada69899 Rice gene	200	14	70.0	3089	2	AAQ51188	Aaq51188 Human GFA
c 128	14	70.0	1279	2	AAV27347	Aav27347 Streptoco	201	14	70.0	3136	8	ADA05761	Ada05761 Human NOV
c 129	14	70.0	1279	6	AQB84815	Aqb84815 S. pneumo	202	14	70.0	3136	12	ADNG62925	Adng62925 Human NOV
c 130	14	70.0	1279	10	ADC45128	Adc45128 S. pneumo	203	14	70.0	3140	14	ADY16380	Ady16380 DNA encod
c 131	14	70.0	1281	10	ABX06032	Abx06032 S. pneumo	204	14	70.0	3140	14	ADL03974	Adl03974 DNA encod
c 132	14	70.0	1284	2	AAK60295	Aak60295 DNA encod	c 205	14	70.0	4169	4	ABL28598	AbL28598 Drosophil
c 133	14	70.0	1284	13	ADK44764	Adk44764 Streptoco	c 206	14	70.0	4539	3	AAH21113	Aah21113 C. glutam
c 134	14	70.0	1308	13	ADR93018	Adr93018 Novel S.	207	14	70.0	4963	5	AAH78396	Aah78396 DNA encod
c 135	14	70.0	1308	14	AEA56888	Aea56888 Streptoco	208	14	70.0	5030	4	ABL14824	AbL14824 Drosophil
c 136	14	70.0	1375	5	AA573092	Aas73092 DNA encod	209	14	70.0	5358	4	ABL10596	AbL10596 Drosophil
c 137	14	70.0	1448	11	ADP65415	Adp65415 Human ela	210	14	70.0	5403	2	AAV84140	Aav84140 Mouse pro
c 138	14	70.0	1449	6	ABK35552	Abk35552 Gene Elas	211	14	70.0	5502	6	AAV84140	Aav84140 Mouse pro
c 139	14	70.0	1449	12	ADQ20799	Adq20799 Human sof	c 212	14	70.0	5555	5	AA94025	Aas94025 DNA encod
c 140	14	70.0	1492	13	ADR60296	Adr60296 Cotton cd	213	14	70.0	8002	4	ABL10738	AbL10738 Drosophil
141	14	70.0	1522	12	ADP04424	Adp04424 Sea squir	c 214	14	70.0	14283	4	ABL11798	AbL11798 Drosophil
142	14	70.0	1588	8	ACA24312	Aca24312 Prokaryot	c 215	14	70.0	16995	2	AAV52215	Aav52215 Streptoco
c 143	14	70.0	1599	13	ADX65230	Adx65230 Plant ful	c 216	14	70.0	17069	3	AAA35063	Aaa35063 Human ade
c 144	14	70.0	1614	5	AA903035	Aas90305 DNA encod	217	14	70.0	17069	3	AAF21185	Aaf21185 Human low
c 145	14	70.0	1992	5	AA569364	Aas69364 DNA encod	218	14	70.0	17069	10	ABZ96879	Abz96879 Human nuc
146	14	70.0	2046	14	ADW23828	Adw23828 Novel hum	219	14	70.0	17069	11	ABD20728	Abd20728 Human pul
147	14	70.0	2064	14	ADW23834	Adw23834 Novel hum	220	14	70.0	24526	6	AAAD43979	Aad43979 Human tra
148	14	70.0	2097	6	ABA93852	AbA93852 Human GFA	221	14	70.0	24526	13	ADU48434	Adu48434 Human tra
149	14	70.0	2100	14	ADW23832	Adw23832 Novel hum	222	14	70.0	50927	13	ABD32585	Abd32585 Human can
150	14	70.0	2118	14	ADW23838	Adw23838 Novel hum	c 223	14	70.0	80000	12	ADP49338	Adp49338 Human B-c
151	14	70.0	2129	8	ABZ79986	Abz79986 Human NF-	c 224	14	70.0	89328	6	ABL61995	AbL61995 Colon ade
c 152	14	70.0	2170	10	ADK67023	Adk67023 Gene #113	225	14	70.0	93390	10	ADD71350	Add71350 Glutamine
c 153	14	70.0	2197	6	ABS73292	Abs73292 DNA encod	226	14	70.0	96596	10	ADB95968	Ade95968 Human NFA
c 154	14	70.0	2197	6	ABS73293	Abs73293 DNA encod	227	14	70.0	96597	9	ADA02720	Ada02720 Human NFA
c 155	14	70.0	2197	13	ADR25367	Adr25367 Breast ca	228	14	70.0	96597	10	AD872458	Adb72458 Human NFA
156	14	70.0	2227	12	ADQ24524	Adq24524 Human sof	c 229	14	70.0	104729	12	ADQ18615	Adq18615 Human sof
c 157	14	70.0	2306	4	ABL21350	AbL21350 Drosophil	c 230	14	70.0	110000	10	ABS56454_02	Continuation (3 of
c 158	14	70.0	2360	12	ADK70354	Adk70354 Respirato	c 231	14	70.0	137908	11	ADP65634	Adp65634 Human seq
c 159	14	70.0	2393	4	AAH99742	Aah99742 Human pro	c 232	14	70.0	201239	8	ACA64924	Aca64924 Human PLZ
c 160	14	70.0	2406	5	ABV29515	Abv29515 Human pro	c 233	14	70.0	201239	8	ACA64924	Aca64924 Human PLZ
c 161	14	70.0	2407	5	AA587932	Aas87932 DNA encod	c 234	13	65.0	20	12	ADP76645	Adp76645 Chimeric
c 162	14	70.0	2407	5	ABV27136	Abv27136 Human pro	c 235	13	65.0	20	12	ADP77259	Adp77259 Chimeric
c 163	14	70.0	2407	5	ABV21317	Abv21317 Human pro	c 236	13	65.0	24	2	AAV48170	Aav48170 Nucleotid
c 164	14	70.0	2407	5	ABV23653	Abv23653 Human pro	c 237	13	65.0	28	10	ADC40569	Adc40569 Glucosami
c 165	14	70.0	2439	9	ADA02722	Ada02722 Human NFA	c 238	13	65.0	60	6	ABN38974	Abn38974 Human spl
										100	8	ACD78689	Act78689 E. coli K









677	12	60.0	214	10	ADG41500	Adg41500 Human res	750	12	60.0	380	5	ABV15136	Abv15136 Human pro
678	12	60.0	214	10	ADG41499	Adg41499 Human res	751	12	60.0	380	12	ADP94984	Adp94984 Cotton ex
679	12	60.0	214	11	ADI97273	Adi97273 Human res	752	12	60.0	381	13	ACF87055	Acf87055 Human SIR
680	12	60.0	214	11	ADI97274	Adi97274 Human res	753	12	60.0	384	6	ABN25213	Abn25213 Human ORF
681	12	60.0	216	8	ACC55311	Acc55311 Rice endo	754	12	60.0	387	8	ABX52954	Abx52954 Bovine ES
682	12	60.0	217	12	ADF08364	Adf08364 Subacute	755	12	60.0	387	8	ACA49501	Ac49501 Prokaryot
683	12	60.0	217	12	ADF08372	Adf08372 Subacute	756	12	60.0	387	8	ACA35276	Ac35276 Prokaryot
684	12	60.0	221	8	ACA13642	Ac13642 Prokaryot	757	12	60.0	387	8	ACA32075	Ac32075 Prokaryot
685	12	60.0	234	8	ACA14953	Ac14953 Prokaryot	758	12	60.0	388	4	AAI86483	Aai86483 Human pol
686	12	60.0	239	4	AAI68589	Aai68589 A. thalia	759	12	60.0	389	4	AAI81709	Aai81709 Human pol
687	12	60.0	245	3	RAC05871	Rac05871 Human sec	760	12	60.0	390	4	AAS56361	Aas56361 Salmoneil
688	12	60.0	245	4	AAK68067	Aak68067 Corynebac	761	12	60.0	390	4	AAS52602	Aas52602 E. coli D
689	12	60.0	245	4	AAK79507	Aak79507 Human imm	762	12	60.0	390	4	AAH84620	Aah84620 E. coli g
690	12	60.0	246	4	AAK79508	Aak79508 Human imm	763	12	60.0	390	4	AAH84588	Aah84588 E. coli g
691	12	60.0	247	3	RAC20072	Rac20072 Human sec	764	12	60.0	390	5	AAH81470	Aah81470 Escherich
692	12	60.0	253	3	RAC21861	Rac21861 Human sec	765	12	60.0	390	8	ACA54229	Ac54229 Prokaryot
693	12	60.0	265	10	ABX85173	Abx85173 Corn ear-	766	12	60.0	390	8	ACA51680	Ac51680 Prokaryot
694	12	60.0	267	4	AAK61163	Aak61163 Human imm	767	12	60.0	390	8	ACA18721	Ac18721 Prokaryot
695	12	60.0	280	8	ACA14585	Ac14585 Prokaryot	768	12	60.0	391	4	AAI89269	Aai89269 Human pol
696	12	60.0	280	12	ADP93793	Adp93793 Cotton ex	769	12	60.0	391	6	ABL37254	Ab137254 Human col
697	12	60.0	281	12	ACH83377	Ach83377 Human gen	770	12	60.0	393	11	ACH97969	Ach97969 Klebsiell
698	12	60.0	283	5	ABV49985	Abv49985 Human pro	771	12	60.0	393	6	ABS97602	Ab97602 Human epo
699	12	60.0	287	6	ABN23203	Abn23203 Human ORF	772	12	60.0	395	4	AAAL34999	Aal34999 Human mus
700	12	60.0	291	2	AAT90489	Aat90489 Ferredoxi	773	12	60.0	395	5	ABV10068	Abv10068 Human pro
701	12	60.0	296	6	ABL73316	Ab173316 Corn tass	774	12	60.0	395	8	ABX57987	Abx57987 cDNA enco
702	12	60.0	300	3	AAAO0769	Aao0769 Human col	775	12	60.0	395	9	ACH31164	Ach31164 Human bon
703	12	60.0	300	4	AAAO6005	Aao6005 Human rep	776	12	60.0	395	12	ADJ27714	Adj27714 Human mus
704	12	60.0	300	4	ABL98570	Ab198570 Human tes	777	12	60.0	395	12	ADP92728	Adp92728 Cotton ex
705	12	60.0	302	13	ACN57337	Acn57337 Cotton gy	778	12	60.0	398	4	AAK63960	Aak63960 Human imm
706	12	60.0	303	12	ADJ03139	Adj03139 DNA enco	779	12	60.0	398	4	AAI85370	Aai85370 Human pol
707	12	60.0	304	6	ABN75795	Abn75795 Human ORF	780	12	60.0	398	6	ABK64159	Abk64159 Human ben
708	12	60.0	305	4	AAK62665	Aak62665 Human imm	781	12	60.0	399	8	ACA14813	Aca14813 Prokaryot
709	12	60.0	306	11	ACL29388	Acl29388 Rice abio	782	12	60.0	399	11	ACL33880	Acl33880 Rice abio
710	12	60.0	306	13	ACN57425	Acn57425 Cotton gy	783	12	60.0	400	4	AAH69460	Aah69460 Human cer
711	12	60.0	315	4	AAI87231	Aai87231 Human pol	784	12	60.0	402	5	AAH78481	Aah78481 DNA enco
712	12	60.0	316	4	AAK65706	Aak65706 Human imm	785	12	60.0	405	4	AAH70795	Aah70795 Human cer
713	12	60.0	316	4	AAK65705	Aak65705 Human imm	786	12	60.0	406	4	AAI85420	Aai85420 Human pol
714	12	60.0	318	3	AAA42115	Aaa42115 Human sec	787	12	60.0	406	8	ACA15328	Aca15328 Prokaryot
715	12	60.0	318	3	AAI84344	Aai84344 Human pol	788	12	60.0	406	8	AAI89973	Aai89973 Human pol
716	12	60.0	319	4	AAI84344	Aai84344 Human pol	789	12	60.0	410	4	AAI85344	Aai85344 Human pol
717	12	60.0	333	3	AAZ53273	Aaz53273 Neisseria	790	12	60.0	414	4	AAK64095	Aak64095 Human imm
718	12	60.0	336	3	AAZ53274	Aaz53274 Neisseria	791	12	60.0	418	4	AAK64095	Aak64095 Human imm
719	12	60.0	341	4	AAK60871	Aak60871 Human imm	792	12	60.0	419	4	AAI81000	Aai81000 Human pol
720	12	60.0	343	3	AAF11224	Aaf11224 Fusarium	793	12	60.0	420	12	ADQ19997	Adq19997 Human sof
721	12	60.0	343	13	ADU55265	Adu55265 Fusarium	794	12	60.0	420	12	ADQ03163	Adq03163 Arabidops
722	12	60.0	343	14	ADP93268	Adp93268 Fusarium	795	12	60.0	421	5	ABV17982	Abv17982 Human pro
723	12	60.0	345	4	AAH72341	Aah72341 Human cer	796	12	60.0	422	5	AAH81899	Aas81899 DNA enco
724	12	60.0	346	4	AAI90751	Aai90751 Human pol	797	12	60.0	423	4	AAK71921	Aak71921 Human imm
725	12	60.0	346	4	AAI91700	Aai91700 Human pol	798	12	60.0	423	6	ABK79803	Abk79803 Bacillus
726	12	60.0	349	4	AAK61279	Aak61279 Human imm	799	12	60.0	423	9	ADA32293	Ada32293 DNA enco
727	12	60.0	350	10	ABX99144	Abx99144 Rice endo	800	12	60.0	424	4	AAI85344	Aai85344 Human pol
728	12	60.0	353	4	AAAS8884	Aas8884 Human can	801	12	60.0	425	4	AAK58090	Aak58090 Human imm
729	12	60.0	354	5	ABV53871	Abv53871 Human pro	802	12	60.0	434	13	ACF91475	Acf91475 Human SIR
730	12	60.0	354	14	ADY66713	Ady66713 S. mansoni	803	12	60.0	435	13	ACF89166	Acf89166 Human SIR
731	12	60.0	355	4	AAK58942	Aak58942 Human imm	804	12	60.0	436	4	AAK56279	Aak56279 Human imm
732	12	60.0	360	3	AAK30892	Aak30892 Breast ca	805	12	60.0	436	4	AAK78778	Aak78778 Human imm
733	12	60.0	360	4	AAK74118	Aak74118 Human imm	806	12	60.0	438	8	ACA14646	Aca14646 Prokaryot
734	12	60.0	361	3	RAC28849	Rac28849 Human sec	807	12	60.0	438	8	ACA14480	Ac14480 Prokaryot
735	12	60.0	363	4	AAAL22937	Aal22937 Human bre	808	12	60.0	439	4	AAK55333	Aak55333 Human imm
736	12	60.0	363	4	AAAL14075	Aal14075 Human bre	809	12	60.0	439	5	AAF93324	Aaf93324 Small int
737	12	60.0	364	4	AAI91421	Aai91421 Human pol	810	12	60.0	439	8	ABX53746	Abx53746 Bovine ES
738	12	60.0	364	10	ACD95033	Ac95033 Human col	811	12	60.0	440	4	AAK89024	Aak89024 Human dig
739	12	60.0	365	3	RAC29635	Rac29635 Human sec	812	12	60.0	445	4	AAAL25089	Aal25089 Human bre
740	12	60.0	366	4	AAAL24245	Aal24245 Human bre	813	12	60.0	445	4	AAH72933	Aah72933 Human cer
741	12	60.0	368	4	AAI89550	Aai89550 Human pol	814	12	60.0	447	10	ACF71485	Acf71485 Phototrab
742	12	60.0	370	4	AAAL33598	Aal33598 Human bre	815	12	60.0	447	13	AD57714	Ad57714 Bacterial
743	12	60.0	374	6	ABA91568	Ab91568 Expressed	816	12	60.0	447	13	ADX63517	Adx63517 Plant ful
744	12	60.0	374	10	ABX93676	Abx93676 Expressed	817	12	60.0	451	6	ABL36955	Ab136955 Human col
745	12	60.0	374	14	ADY86165	Ady86165 Human EG-	818	12	60.0	459	5	ABV33025	Abv33025 Human pro
746	12	60.0	376	8	ACA14924	Ac14924 Prokaryot	819	12	60.0	460	4	AAK82953	Aak82953 Human imm
747	12	60.0	378	4	AAI80786	Aai80786 Human pol	820	12	60.0	461	7	ADS72594	Ad72594 Human kid
748	12	60.0	378	10	ADBE1417	Adbe1417 Rat gene	821	12	60.0	461	7	ADW41448	Adw41448 cDNA elev
749	12	60.0	378	10	ADP46053	Adp46053 Rat gene	822	12	60.0	467	4	AAAS34898	Aas34898 cDNA enco



```
c 969 12 60.0 608 6 ABN64400 Human can
970 12 60.0 610 11 ACN91572 Breast ca
971 12 60.0 612 5 ABV41950 Human pro
c 972 12 60.0 612 6 ABK30448 Human G-p
973 12 60.0 614 5 ABV33573 Human pro
974 12 60.0 614 5 ABV42493 Human pro
c 975 12 60.0 615 3 AAL16210 Human col
c 976 12 60.0 616 13 ADR60466 Cotton cd
c 977 12 60.0 617 13 ACN57656 Cotton gy
c 978 12 60.0 618 4 AAS42855 Human G P
c 979 12 60.0 625 8 ACA13938 Prokaryot
c 980 12 60.0 625 13 ADQ57160 Novel can
981 12 60.0 626 10 ADD16075 cDNA (Seq
c 982 12 60.0 628 13 ADX48103 Plant ful
983 12 60.0 629 13 ADQ54385 Novel can
c 984 12 60.0 629 13 ADX09341 Plant ful
c 985 12 60.0 630 3 AAC44029 Zea mays
986 12 60.0 630 4 AAS27145 cDNA enco
987 12 60.0 630 4 AAS30220 DNA enco
988 12 60.0 630 6 ABN60052 Human can
989 12 60.0 630 8 ACD01475 Human pol
c 990 12 60.0 630 10 ADB93323 Human CDN
c 991 12 60.0 633 11 ACN86193 Breast ca
992 12 60.0 634 10 ADE09829 Novel DNA
c 993 12 60.0 634 13 ADX11530 Plant ful
994 12 60.0 635 5 ADL37561 Human ova
995 12 60.0 635 5 ADI72418 Human ova
996 12 60.0 636 3 AAF13811 Aspergill
997 12 60.0 636 4 AAH52675 S. epider
998 12 60.0 636 13 ADU57852 Aspergill
999 12 60.0 636 14 ADZ95855 Aspergill
c1000 12 60.0 637 13 ADX32579 Plant ful
```

## ALIGNMENTS

```
RESULT 1
ADN97882
ID ADN97882 standard; DNA; 20 BP.
XX
AC ADN97882;
XX
DT 01-JUL-2004 (first entry)
XX
DE Monkey foxhead box O1a sequence inhibitory oligo #1.
XX
KW SS; cytostatic; antidiabetic; foxhead box O1a inhibitor;
KW foxhead box O1a; hyperproliferative disorder; cancer; rhabdomyosarcoma;
KW diabetes; H-ras gene; antisense; gene expression; primer.
XX
OS Synthetic.
XX
PN W02004031350-A2.
XX
PD 15-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030352.
XX
PR 26-SEP-2002; 2002US-00260203.
XX
PS (AMGE-) AMGEN INC.
PA (ISIS-) ISIS PHARM INC.
XX
PI Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;
XX WPI; 2004-330164/30.
XX
PT New compounds, particularly antisense oligonucleotides, targeted to a
PT nucleic acid molecule encoding foxhead box O1a, useful for treating
PT cancer, or type 2 diabetes.
XX
PS Example 25; SEQ ID NO 172; 146pp; English.
```

```
XX The invention relates to a compound 8-80 nucleobases in length targeted
CC to a nucleic acid molecule encoding foxhead box O1a, where the compound
CC is at least 70% complementary to a nucleic acid molecule encoding
CC foxhead box O1a and modulates expression of foxhead box O1a by at least
CC 10%. The compound is useful for treating an animal having a disease or
CC condition associated with foxhead box O1a, e.g. a hyperproliferative
CC disorder (cancer, preferably rhabdomyosarcoma), or type 2 diabetes. This
CC sequence corresponds to an oligonucleotide targeted to the monkey foxhead
CC box O1a genes in order to inhibit gene expression.
XX
SQ Sequence 20 BP; 4 A; 4 C; 6 G; 6 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCTTGGTTGGCAACACAT 20
Db 1 GCTTGGTTGGCAACACAT 20
RESULT 2
ADN40818
ID ADN40818 standard; DNA; 20 BP.
XX
AC ADN40818;
XX
DT 12-AUG-2004 (first entry)
XX
DE Monkey foxhead box O1a DNA antisense oligonucleotide #1.
XX
KW Monkey; foxhead box O1a; ss; antisense oligonucleotide;
KW phosphorothioate linkage; 2'-O-methoxyethyl sugar moiety;
KW 5-methylcytosine; hyperproliferative disorder; cancer; rhabdomyosarcoma;
KW type 2 diabetes; cytostatic; antidiabetic.
XX
OS Primates.
XX
PN US2004097459-A1.
XX
PD 20-MAY-2004.
XX
PF 25-SEP-2003; 2003US-00671074.
XX
PR 26-SEP-2002; 2002US-00260203.
XX
PA (DOBI/) DOBIE K W.
PA (SHAN/) BHANOT S.
PA (VENI/) VENIANT-ELLISON M.
PA (LIND/) LINDBERG R A.
PA (SHUT/) SHUTTER J R.
XX
PI Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;
XX WPI; 2004-389194/36.
XX
PT New compounds, particularly antisense oligonucleotides, targeted to a
PT nucleic acid molecule encoding foxhead box O1a, useful for treating
PT cancer, or type 2 diabetes.
XX
PS Example 25; SEQ ID NO 172; 80pp; English.
XX
CC The invention relates to a compound targeted to a nucleic acid molecule
CC encoding the human foxhead box O1a polypeptide. The compound is an
CC antisense oligonucleotide that specifically hybridizes with the nucleic
CC acid and inhibits expression of the polypeptide. The antisense
CC oligonucleotide comprises at least one modified internucleoside linkage
CC i.e. a phosphorothioate linkage, at least one modified sugar moiety,
CC preferably a 2'-O-methoxyethyl sugar moiety, or at least one modified
CC nucleobase comprising a 5-methylcytosine. The antisense compounds are
CC useful for modulating the expression of the human foxhead box O1a
CC polypeptide and in preparation of a composition for treating
```

CC hyperproliferative disorders, e.g. cancer, preferably rhabdomyosarcoma,  
CC and type 2 diabetes. This sequence represents an antisense  
CC oligonucleotide targeted to DNA encoding the monkey forkhead O1A  
CC polypeptide of the invention.

XX  
SQ Sequence 20 BP; 4 A; 4 C; 6 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.1; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0;

Qy 1 GCTTTGGTTGGCACAACAT 20  
Db 1 GCTTTGGTTGGCACAACAT 20

## RESULT 3

AAS14337/c  
ID AAS14337 standard; cDNA; 254 BP.

XX  
AC AAS14337;

XX  
DT 27-FEB-2002 (first entry)

XX  
DE Human endometriosis-related marker DD12.

XX  
KW Human; endometriosis-related marker; endometrial cell; DD12;  
XX differential display; antiinflammatory; ss.

XX  
OS Homo sapiens.

XX  
PN WO200162959-A2.

XX  
PD 30-AUG-2001.

XX  
PF 26-FEB-2001; 2001WO-CA000245.

XX  
PR 25-FEB-2000; 2000US-0185063P.

XX  
PR 17-AUG-2000; 2000US-0225745P.

XX  
PA (PROC-) PROCREA BIOSCIENCES INC.

XX  
PI Baban S, Bernard M, Cherry B, Gosselin D, Hugo P, Malette B;  
XX Miron P, Prive C, Shazand K;

XX  
PI WPI; 2001-648220/74.

XX  
PT Determining likelihood of endometriosis in female subject, involves  
XX obtaining endometrial cell sample from the subject, and assaying the  
XX cells for the expression level of at least one endometriosis-related  
XX marker.

XX  
PS Claim 44; Page 56; 92pp; English.

XX  
CC The present invention relates to a method of determining the likelihood  
XX of endometriosis in a female subject. The method comprises obtaining a  
XX sample of endometrial cells from a female subject, and assaying for the  
XX expression level of at least one endometriosis-related marker. The  
XX endometriosis-related markers are identified by differential display and  
XX are referred to as DD1-DD16. The markers are differentially expressed in  
XX the endometrial cells of females with endometriosis compared to  
XX endometriosis-free females. The method is useful as a diagnostic method  
XX for the detection of endometriosis in a female subject. The method is  
XX more rapid, non-invasive, much less complicated, and much less costly  
XX than laparoscopy. In contrast to the currently-available methods, the  
XX method is suitable for direct measurement of expression levels of  
XX endometriosis-related genes that are expressed differentially in  
XX endometrial cells depending on the presence/absence of endometriosis and  
XX the stage of the disease with relatively high levels of sensitivity and  
XX specificity. AAS14326-AAS14341 represent the endometriosis-related  
XX markers (DD1-DD16) of the present invention

XX  
SQ Sequence 254 BP; 77 A; 52 C; 57 G; 68 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 5; Length 254;  
Best Local Similarity 100.0%; Pred. No. 0.096;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCACAACAT 20  
Db 90 GCTTTGGTTGGCACAACAT 71

## RESULT 4

ADK61538  
ID ADK61538 standard; DNA; 520 BP.

XX  
AC ADK61538;

XX  
DT 06-MAY-2004 (first entry)

XX  
DE Ovarian cancer-related DNA #693 with altered ovarian cancer expression.

XX  
KW ds; gene; ovarian tumor; BRCA-1-like; BRCA-2-like; non-BRCA-like;  
XX gene expression; primer; cancer.

XX  
OS Homo sapiens.

XX  
PN WO2003068054-A2.

XX  
PD 21-AUG-2003.

XX  
PF 13-FEB-2003; 2003WO-US004688.

XX  
PR 13-FEB-2002; 2002US-0357031P.

XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX  
PA (SLOK ) SLOAN KETTERING INST CANCER RES.

XX  
PI Jazaeri AA, Boyd J, Liu ET;

XX  
PI WPI; 2003-689589/65.

XX  
PT Classifying an ovarian tumor as a BRCA-1-like or BRCA-2-like or non-BRCA-  
XX like tumor by determining a pattern of expression in the ovarian tumor of  
XX several markers.

XX  
PS Disclosure; SEQ ID NO 708; 137pp; English.

XX  
CC The invention relates to a method of classifying an ovarian tumor as a  
XX BRCA-1-like or BRCA-2-like or non-BRCA-like tumor by: (1) determining a  
XX pattern of expression in the ovarian tumor of several markers given in  
XX the specification; and (2) comparing a similarity of the pattern of  
XX expression of the markers in the ovarian tumor to a pattern of expression  
XX of the markers in a comparison tissue of a known BRCA-1-like or BRCA-2-  
XX like or non-BRCA-like tumor. The method is useful for classifying an  
XX ovarian tumor as a BRCA-1-like or BRCA-2-like or non-BRCA-like tumor.  
XX This sequence corresponds to an ovarian cancer -related gene having an  
XX altered pattern of expression in ovarian cancer. (Note: The sequence data  
XX for this patent did not form part of the printed specification but was  
XX obtained in electronic format from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences).

XX  
SQ Sequence 520 BP; 119 A; 122 C; 135 G; 134 T; 0 U; 10 Other;

Query Match 100.0%; Score 20; DB 10; Length 520;  
Best Local Similarity 100.0%; Pred. No. 0.095;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCACAACAT 20  
Db 163 GCTTTGGTTGGCACAACAT 182

## RESULT 5

ADF79645/c

```
ID ADF79645 standard; DNA; 578 BP.
XX AC
XX ADF79645;
XX DT
XX 26-FEB-2004 (first entry)
XX DE
XX Leukaemia-related DNA sequence #201.
XX KW
XX Cytostatic; Gene therapy; leukaemia; ss.
XX OS
XX Unidentified.
XX PN
XX WO2003039443-A2.
XX PD
XX 15-MAY-2003.
XX PF
XX 04-NOV-2002; 2002WO-EP012303.
XX PR
XX 05-NOV-2001; 2001EP-00126244.
XX PR
XX 30-APR-2002; 2002EP-00009758.
XX PA
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PA
XX (UJLU-) UNIV LUDWIG MAXIMILIANS.
XX PA
XX (HAFE/) HAFERLACH T.
XX PA
XX (SCHO/) SCHOCH C.
XX PA
XX (KERN/) KERN W.
XX PI
XX Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;
XX PI
XX Eils R, Brors B, Mergenthaler S;
XX DR
XX WPI; 2003-505037/47.
XX PT
XX Determining the subtype of leukemia cells and whether a patient sample
XX PT
XX contains leukemia cells or other cells, useful for treating leukemia,
XX PT
XX comprises determining the expression profile of a group of markers in a
XX PT
XX patient sample.
XX PS
XX Disclosure; SEQ ID NO 201; 2938pp; English.
XX CC
XX The present invention relates to a method (M1) for determining the
XX CC
XX subtype of leukaemia cells and whether a patient sample contains
XX CC
XX leukaemia cells. The method comprises determining the expression profile
XX CC
XX of a group of markers in a patient sample. The method is useful for
XX CC
XX determining the presence of leukaemia cells, its types or subtypes, and
XX CC
XX for the preparation of a medicament for treating leukaemia.
XX SQ
XX Sequence 578 BP; 153 A; 164 C; 137 G; 124 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 10; Length 578;
XX Best Local Similarity 100.0%; Pred. No. 0.095;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCTTTGGTTGGGCAACACAT 20
XX DB
XX 437 GCTTTGGTTGGGCAACACAT 418
XX
XX RESULT 6
XX AAS85052/c
XX ID
XX AAS85052 standard; cDNA; 2514 BP.
XX AC
XX AAS85052;
XX XX
XX 13-FEB-2002 (first entry)
XX DT
XX
XX DE
XX DNA encoding novel human diagnostic protein #20856.
XX KW
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX XX
XX Homo sapiens.
XX OS
XX WO200175067-A2.
XX PN
XX
XX ADF79645 standard; DNA; 578 BP.
XX AC
XX ADF79645;
XX DT
XX 26-FEB-2004 (first entry)
XX DE
XX Leukaemia-related DNA sequence #201.
XX KW
XX Cytostatic; Gene therapy; leukaemia; ss.
XX OS
XX Unidentified.
XX PN
XX WO2003039443-A2.
XX PD
XX 15-MAY-2003.
XX PF
XX 04-NOV-2002; 2002WO-EP012303.
XX PR
XX 05-NOV-2001; 2001EP-00126244.
XX PR
XX 30-APR-2002; 2002EP-00009758.
XX PA
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PA
XX (UJLU-) UNIV LUDWIG MAXIMILIANS.
XX PA
XX (HAFE/) HAFERLACH T.
XX PA
XX (SCHO/) SCHOCH C.
XX PA
XX (KERN/) KERN W.
XX PI
XX Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;
XX PI
XX Eils R, Brors B, Mergenthaler S;
XX DR
XX WPI; 2003-505037/47.
XX PT
XX Determining the subtype of leukemia cells and whether a patient sample
XX PT
XX contains leukemia cells or other cells, useful for treating leukemia,
XX PT
XX comprises determining the expression profile of a group of markers in a
XX PT
XX patient sample.
XX PS
XX Disclosure; SEQ ID NO 201; 2938pp; English.
XX CC
XX The present invention relates to a method (M1) for determining the
XX CC
XX subtype of leukaemia cells and whether a patient sample contains
XX CC
XX leukaemia cells. The method comprises determining the expression profile
XX CC
XX of a group of markers in a patient sample. The method is useful for
XX CC
XX determining the presence of leukaemia cells, its types or subtypes, and
XX CC
XX for the preparation of a medicament for treating leukaemia.
XX SQ
XX Sequence 578 BP; 153 A; 164 C; 137 G; 124 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 10; Length 578;
XX Best Local Similarity 100.0%; Pred. No. 0.095;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCTTTGGTTGGGCAACACAT 20
XX DB
XX 437 GCTTTGGTTGGGCAACACAT 418
XX
XX RESULT 6
XX AAS85052/c
XX ID
XX AAS85052 standard; cDNA; 2514 BP.
XX AC
XX AAS85052;
XX XX
XX 13-FEB-2002 (first entry)
XX DT
XX
XX DE
XX DNA encoding novel human diagnostic protein #20856.
XX KW
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX XX
XX Homo sapiens.
XX OS
XX WO200175067-A2.
XX PN
XX
XX 11-OCT-2001.
XX PD
XX
XX 30-MAR-2001; 2001WO-US008631.
XX PF
XX
XX 31-MAR-2000; 2000US-00540217.
XX PR
XX 23-AUG-2000; 2000US-00649167.
XX PR
XX
XX (HYSE-) HYSEQ INC.
XX PA
XX
XX Drmanac RT, Liu C, Tang YT;
XX PI
XX WPI; 2001-639362/73.
XX DR
XX P-PSDB; ABG20865.
XX DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT
XX diagnostics, forensics, gene mapping, identification of mutations
XX PT
XX responsible for genetic disorders or other traits and to assess
XX PT
XX biodiversity.
XX PR
XX
XX Claim 1; SEQ ID NO 20856; 103pp; English.
XX PS
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC
XX and in recombinant production of (II). The polynucleotides are also used
XX CC
XX in diagnostics as expressed sequence tags for identifying expressed
XX CC
XX genes. (I) is useful in gene therapy techniques to restore normal
XX CC
XX activity of (II) or to treat disease states involving (II). (II) is
XX CC
XX useful for generating antibodies against it, detecting or quantitating a
XX CC
XX polypeptide in tissue, as molecular weight markers and as a food
XX CC
XX supplement. (II) and its binding partners are useful in medical imaging
XX CC
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC
XX involving aberrant protein expression or biological actions in
XX CC
XX polypeptide and polynucleotide sequences have applications in
XX CC
XX diagnostics, forensics, gene mapping, identification of mutations
XX CC
XX and to produce other types of data and products dependent on DNA and
XX CC
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC
XX coding sequences of the invention. Note: The sequence data for this
XX CC
XX patent did not appear in the printed specification, but was obtained in
XX CC
XX electronic format directly from WIPO at
XX CC
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ
XX Sequence 2514 BP; 658 A; 738 C; 620 G; 498 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 5; Length 2514;
XX Best Local Similarity 100.0%; Pred. No. 0.093;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCTTTGGTTGGGCAACACAT 20
XX DB
XX 2466 GCTTTGGTTGGGCAACACAT 2447
XX
XX RESULT 7
XX ACC72829/c
XX ID
XX ACC72829 standard; cDNA; 2827 BP.
XX XX
XX ACC72829;
XX AC
XX
XX 09-JUL-2003 (first entry)
XX DT
XX
XX DE
XX Human cancer related protein encoding cDNA SEQ ID NO:167.
XX XX
XX Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
XX KW
XX heart disease; atherosclerosis; endometriosis; gene; ss.
XX KW
XX
XX Homo sapiens.
XX OS
XX WO2003025138-A2.
XX XX
XX 27-MAR-2003.
XX PD
```



XX 17-SEP-2002; 2002WO-US029560.  
 XX 17-SEP-2001; 2001US-0323469P.  
 PR 20-SEP-2001; 2001US-0323887P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 08-FEB-2002; 2002US-0355145P.  
 PR 08-FEB-2002; 2002US-0355257P.  
 PR 12-APR-2002; 2002US-0372246P.  
 XX (EOSB-) EOS BIOTECHNOLOGY INC.  
 PA Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KB;  
 PI Zlotnik A;  
 XX WPI; 2003-354600/33.  
 DR P-PSDB; ABR58678.  
 XX New genes that are up-regulated or down-regulated in cancers, useful as  
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as  
 PT therapeutic targets for screening drugs for treating these diseases.  
 XX  
 PS Claim 8; Page 718-719; 767pp; English.  
 XX  
 CC The present invention describes an isolated nucleic acid molecule, which  
 CC comprises the sequence of any of the genes that are up-regulated or down-  
 CC regulated in specific cancers (e.g. about 1031 genes up-regulated in  
 CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer  
 CC related gene nucleotide sequences which encode the proteins given in  
 CC ABR58521 to ABR59709. Also described: (1) determining the presence or  
 CC absence of a pathological cell in a patient; (2) an expression vector  
 CC comprising a nucleic acid molecule described above; (3) a host cell  
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by  
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide  
 CC of (4); (6) specifically targeting a compound to a pathological cell in a  
 CC patient by administering to the patient the antibody above; and (7) a  
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or  
 CC therapeutic targets. In particular, the nucleic acid is useful for  
 CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,  
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,  
 CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,  
 CC atherosclerosis and endometriosis. The nucleic acid is also useful in  
 CC drug screening, particularly for identifying agents for treating these  
 CC pathologies  
 XX  
 SQ Sequence 2827 BP; 732 A; 831 C; 693 G; 571 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 10; Length 2827;  
 Best Local Similarity 100.0%; Pred. No. 0.093;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GCTTTGGTTGGGCAACACAT 20  
 |||||  
 Db 2647 GCTTTGGTTGGGCAACACAT 2628  
 RESULT 8  
 ADQ03121/c  
 ID ADQ03121 standard; DNA; 2827 BP.  
 XX  
 AC ADQ03121;  
 XX  
 DT 26-AUG-2004 (first entry)  
 XX  
 XX PAX3-FKHR fusion gene translocation DNA seqid 1.  
 DE  
 XX fusion transcript identification; chromosome abnormality;  
 KW genetic disorder; degenerative disease; cancer; leukaemia;  
 KW chronic myelogenous leukaemia; CML; gene rearrangement; PAX3-FKHR;  
 KW translocation; Alveolar rhabdomyosarcoma; ARMS; BS;  
 XX expressed sequence tag; EST.  
 KW  
 OS Unidentified.

XX US2004110227-A1.  
 XX 10-JUN-2004.  
 XX 19-MAR-2003; 2003US-00391530.  
 XX 19-MAR-2002; 2002US-0365076P.  
 XX (LEVA/) LEVANON E.  
 PA (TOPO/) TOPORIK A.  
 PA (AKIV/) AKIVA P.  
 XX Levanon E, Toporik A, Akiva P;  
 PI WPI; 2004-532613/51.  
 XX  
 DR Identifying putative fusion transcripts, by aligning first database of  
 PT annotated polynucleotide with database of expressed polynucleotide, and  
 PT identifying expressed polynucleotide complementary to non-contiguous  
 PT sequences of first database.  
 XX  
 PS Example 9; SEQ ID NO 1; 133pp; English.  
 XX  
 CC The invention describes a method of identifying (M1) putative fusion  
 CC transcripts. The method involves computationally aligning a first  
 CC database of annotated polynucleotide sequences with a second database of  
 CC expressed polynucleotide sequences, and identifying in the second  
 CC database an expressed polynucleotide sequence complementary to at least  
 CC two non-contiguous sequences (S1) of the first database, where the  
 CC expressed polynucleotide sequence identified is a putative fusion  
 CC transcript, and (S1) is chosen from non-homologous polynucleotide  
 CC sequences mapped to different chromosomes, polynucleotide sequences  
 CC mapped to different loci of a single chromosome and polynucleotide  
 CC sequences mapped to a single locus and not being a part of a splice  
 CC isoform, the expressed polynucleotide sequence identified is a putative  
 CC fusion transcript. (M1) is useful for identifying putative fusion  
 CC transcripts, for identifying transition points in fusion transcripts, for  
 CC identifying polynucleotide sequence associated with a disorder associated  
 CC with gene rearrangements, for identifying polypeptides resulting from  
 CC putative fusion events, for detecting nucleic acid sequence chimerism  
 CC indicative of predisposition for disorders associated with genetic  
 CC rearrangements, and for identifying putative mutagenic agents. (M1) is  
 CC useful for identifying gene rearrangements, chromosome abnormalities,  
 CC which are associated with genetic disorders, agents known to cause  
 CC degenerative diseases such as cancer e.g., leukaemia, chronic myelogenous  
 CC leukaemia (CML), and in diagnosis of various diseases related to gene  
 CC rearrangements. This sequence represents a PAX3-FKHR t(2;13)(q35;q14)  
 CC translocation polynucleotide associated with Alveolar rhabdomyosarcoma  
 CC (ARMS).  
 XX  
 SQ Sequence 2827 BP; 732 A; 831 C; 693 G; 571 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 12; Length 2827;  
 Best Local Similarity 100.0%; Pred. No. 0.093;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GCTTTGGTTGGGCAACACAT 20  
 |||||  
 Db 2647 GCTTTGGTTGGGCAACACAT 2628  
 RESULT 9  
 ADQ03123/c  
 ID ADQ03123 standard; DNA; 3200 BP.  
 XX  
 AC ADQ03123;  
 XX  
 DT 26-AUG-2004 (first entry)  
 XX  
 XX PAX3-FKHR fusion gene translocation DNA seqid 3.  
 DE  
 XX fusion transcript identification; chromosome abnormality;  
 KW

KW genetic disorder; degenerative disease; cancer; leukaemia;  
KW chronic myelogenous leukaemia; CML; gene rearrangement; PAX3-FKHR;  
KW translocation; Alveolar rhabdomyosarcoma; ARMS; BS;  
KW expressed sequence tag; EST.  
XX Homo sapiens.  
XX US2004110227-A1.  
XX 10-JUN-2004.  
XX 19-MAR-2003; 2003US-00391530.  
XX 19-MAR-2002; 2002US-0365076P.  
XX (LEVA/) LEVANON E.  
XX (TOPO/) TOPORIK A.  
XX (AKIV/) AKIVA P.  
XX Levanon E, Toporik A, Akiva P;  
XX WPI; 2004-533613/51.  
XX Identifying putative fusion transcripts, by aligning first database of  
XX annotated polynucleotide with database of expressed polynucleotide, and  
XX identifying expressed polynucleotide complementary to non-contiguous  
XX sequences of first database.  
XX Example 9; SEQ ID NO 3; 133pp; English.  
XX The invention describes a method of identifying (M1) putative fusion  
XX transcripts. The method involves computationally aligning a first  
XX database of annotated polynucleotide sequences with a second database of  
XX expressed polynucleotide sequences, and identifying in the second  
XX database an expressed polynucleotide sequence complementary to at least  
XX two non-contiguous sequences (S1) of the first database, where the  
XX expressed polynucleotide sequence identified is a putative fusion  
XX transcript, and (S1) is chosen from non-homologous polynucleotide  
XX sequences mapped to different chromosomes. polynucleotide sequences  
XX mapped to different loci of a single chromosome and polynucleotide  
XX sequences mapped to a single locus and not being a part of a splice  
XX isoform, the expressed polynucleotide sequence identified is a putative  
XX fusion transcript. (M1) is useful for identifying putative fusion  
XX transcripts, for identifying transition points in fusion transcripts, for  
XX identifying polynucleotide sequence associated with a disorder associated  
XX with gene rearrangements, for identifying polypeptides resulting from  
XX putative fusion events, for detecting nucleic acid sequence chimerism  
XX indicative of predisposition for disorders associated with genetic  
XX rearrangements, and for identifying putative mutagenic agents. (M1) is  
XX useful for identifying gene rearrangements, chromosome abnormalities,  
XX which are associated with genetic disorders, agents known to cause  
XX degenerative diseases such as cancer e.g., leukaemia, chronic myelogenous  
XX leukaemia (CML), and in diagnosis of various diseases related to gene  
XX rearrangements. This sequence represents a PAX3-FKHR t(2;13)(q35;q14)  
XX translocation polynucleotide associated with Alveolar rhabdomyosarcoma  
XX (ARMS).  
XX Sequence 3200 BP; 857 A; 869 C; 754 G; 720 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 12; Length 3200;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCCTTGGTTGGGCAACACAT 20  
DB 2511 GCCTTGGTTGGGCAACACAT 2492  
RESULT 10  
ADR83528/c  
ID ADR83528 standard; DNA; 3421 BP.  
XX AC ADR83528;

XX 02-DEC-2004 (first entry)  
XX Human fork head domain protein DNA, target gene of miRNA.  
XX human; ds; miRNA; microRNA; ontogenesis; cell therapy; cancer;  
XX immune disease; nerve disorder; amyotrophic lateral sclerosis;  
KW Parkinson's disease; Alzheimer's disease; inflammatory disease;  
KW siRNA silencing precursor; cytosolic; immunosuppressive; nototropic;  
KW neuroprotective; antiinflammatory; immunotherapy;  
KW fork head domain protein.  
XX Homo sapiens.  
XX WO2004076622-A2.  
XX 10-SEP-2004.  
XX 10-FEB-2004; 2004WO-JP001433.  
XX 10-FEB-2003; 2003US-0445829P.  
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX Taira K, Kawasaki H;  
XX WPI; 2004-653393/63.  
XX Modulating expression of a target gene in a cell, for treating cancer, an  
XX immune disease, or a nerve disorder, comprises introducing into the cell  
XX a polynucleotide that forms a duplex region with an mRNA transcribed from  
XX the target gene.  
XX Claim 9; SEQ ID NO 430; 865pp; English.  
XX This invention relates to a novel method for modulating the expression of  
XX a target gene in a cell. Specifically, it refers to the introduction into  
XX a cell of a polynucleotide that forms a duplex region with an mRNA  
XX transcribed from the target gene, where the duplex region comprises a  
XX mammalian miRNA target region i.e. a non-coding microRNA (miRNA) that  
XX regulates mRNA at a post-transcriptional level. The present invention  
XX describes a method for controlling ontogenesis of a mammal, function of a  
XX mammalian cell, differentiation of a mammalian cell or viability of a  
XX mammalian cell in the post-transcriptional phase, which comprises  
XX introducing a plasmid vector comprising a promoter and nucleic acid  
XX molecule expressing an miRNA or siRNA silencing precursor to the miRNA.  
XX Accordingly, it provides a cell therapy method for treating cancer,  
XX immune disease, nerve disorder (e.g. amyotrophic lateral sclerosis,  
XX Parkinson's disease, or Alzheimer's disease) or an inflammatory disease  
XX by introducing into the cell the miRNA, siRNA silencing precursor to the  
XX miRNA or the plasmid vector. As such, they can be developed into  
XX pharmaceutical compositions that exhibit cytostatic, immunosuppressive,  
XX nototropic, neuroprotective and antiinflammatory activities and hence can  
XX be used for immunotherapy. This polynucleotide sequence is a human target  
XX gene whose expression is modulated by miRNAs of the invention.  
XX Sequence 3421 BP; 852 A; 859 C; 805 G; 905 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 13; Length 3421;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCCTTGGTTGGGCAACACAT 20  
DB 1926 GCCTTGGTTGGGCAACACAT 1907  
RESULT 11  
ABS73280/c  
ID ABS73280 standard; DNA; 3517 BP.  
XX AC ABS73280;

DT 04-DEC-2002 (first entry)  
DE DNA encoding human translocation (2; 13) (q35; q14) protein.  
XX Chromosome aberration; oncogenic fusion protein; cancer;  
KW proliferative disease; cellular protein isoform; heat shock protein 90;  
KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;  
KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;  
KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;  
KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;  
KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;  
KW rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.  
XX Homo sapiens.  
OS  
XX  
XX WO200269900-A2.  
FN  
XX  
XX 12-SEP-2002.  
PD  
XX  
XX 01-MAR-2002; 2002WO-US006518.  
PF  
XX  
XX 01-MAR-2001; 2001US-0272751P.  
PR  
XX  
XX (CONF-) CONFORMA THERAPEUTICS CORP.  
PA  
XX  
XX Fritz LC, Burrows RJ;  
PI  
XX  
XX WPI; 2002-698710/75.  
DR  
XX  
XX P-PSDB; ABG95079.  
DR  
XX  
XX Treating genetically-defined disease associated with chromosomal  
PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
PT diseases, involves administering an inhibitor of heat shock protein 90.  
PT  
XX  
XX Disclosure; Page 234-235; 389pp; English.  
PS  
XX  
XX The invention describes a method of treating genetically-defined disease  
CC associated with chromosomal aberrations yielding oncogenic fusion  
CC proteins (I), treating cancerous cells containing (I) in a heterogeneous  
CC cell population, treating proliferative diseases associated with mutant  
CC protein or cellular protein isoforms (II) dependent on heat shock protein  
CC (HSP)-90, or selectively treating cells expressing (II) involving  
CC administering HSP90-inhibitor. The method is useful for treating  
CC genetically-defined disease with chromosomal aberration yielding  
CC oncogenic fusion protein, treating cancerous cells containing fusion  
CC protein in heterogeneous cell population, treating proliferative disease  
CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
CC p53), or selectively treating cells expressing mutant protein or cellular  
CC protein isoform in a patient heterozygous for (II). The method is useful  
CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML,  
CC or a disease characterised by a solid tumour such as papillary thyroid  
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
CC synovial sarcoma. The method is also useful for treating viral  
CC infections. This represents the DNA sequence of a chromosome aberration  
XX  
SQ Sequence 3517 BP; 962 A; 841 C; 756 G; 958 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 6; Length 3517;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCTTTGGTTGGGCAACACAT 20  
|||||  
Db 2022 GCTTTGGTTGGGCAACACAT 2003  
RESULT 12  
ID ADQ03122 standard; DNA; 3517 BP.  
XX  
AC ADQ03122;

XX 26-AUG-2004 (first entry)  
DT  
XX  
XX PAX3-FKHR fusion gene translocation DNA seqid 2.  
DE  
XX  
XX fusion transcript identification; chromosome abnormality;  
KW genetic disorder; degenerative disease; cancer; leukaemia;  
KW chronic myelogenous leukaemia; CML; gene rearrangement; PAX3-FKHR;  
KW translocation; Alveolar rhabdomyosarcoma; ARMS; ss;  
KW expressed sequence tag; EST.  
XX  
XX Homo sapiens.  
OS  
XX  
XX US2004110227-A1.  
FN  
XX  
XX 10-JUN-2004.  
PD  
XX  
XX 19-MAR-2003; 2003US-00391530.  
PF  
XX  
XX 19-MAR-2002; 2002US-0365076P.  
PR  
XX  
XX (LEVA/) LEVANON E.  
PA  
XX (TOPO/) TOPORIK A.  
PA  
XX (AKIV/) AKIVA P.  
PA  
XX  
XX Levanon E, Toporik A, Akiva P;  
PI  
XX  
XX WPI; 2004-532613/51.  
DR  
XX  
XX Identifying putative fusion transcripts, by aligning first database of  
PT annotated polynucleotide with database of expressed polynucleotide, and  
PT identifying expressed polynucleotide complementary to non-contiguous  
PT sequences of first database.  
PT  
XX  
XX Example 9; SEQ ID NO 2; 133pp; English.  
PS  
XX  
XX The invention describes a method of identifying (M1) putative fusion  
CC transcripts. The method involves computationally aligning a first  
CC database of annotated polynucleotide sequences with a second database of  
CC expressed polynucleotide sequences, and identifying in the second  
CC database an expressed polynucleotide sequence complementary to at least  
CC two non-contiguous sequences (S1) of the first database, where the  
CC expressed polynucleotide sequence identified is a putative fusion  
CC transcript, and (S1) is chosen from non-homologous polynucleotide  
CC sequences mapped to different loci of a single chromosome and polynucleotide  
CC mapped to different loci of a single chromosome and polynucleotide  
CC sequences mapped to a single locus and not being a part of a splice  
CC isoform, the expressed polynucleotide sequence identified is a putative  
CC fusion transcript. (M1) is useful for identifying putative fusion  
CC transcripts, for identifying transition points in fusion transcripts, for  
CC identifying polynucleotide sequence associated with a disorder associated  
CC with gene rearrangements, for identifying polypeptides resulting from  
CC putative fusion events, for detecting nucleic acid sequence chimerism  
CC indicative of predisposition for disorders associated with genetic  
CC rearrangements, and for identifying putative mutagenic agents. (M1) is  
CC useful for identifying gene rearrangements, chromosome abnormalities,  
CC which are associated with genetic disorders, agents known to cause  
CC degenerative diseases such as cancer e.g., leukaemia, chronic myelogenous  
CC leukaemia (CML), and in diagnosis of various diseases related to gene  
CC rearrangements. This sequence represents a PAX3-FKHR t(2;13) (q35;q14)  
CC translocation polynucleotide associated with Alveolar rhabdomyosarcoma  
CC (ARMS).  
XX  
SQ Sequence 3517 BP; 962 A; 841 C; 756 G; 958 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 12; Length 3517;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCTTTGGTTGGGCAACACAT 20  
|||||  
Db 2022 GCTTTGGTTGGGCAACACAT 2003

```
RESULT 13
ADN97721/c
ID ADN97721 standard; DNA; 4945 BP.
XX
AC ADN97721;
XX
DT 01-JUL-2004 (first entry)
XX
DE Mouse foxhead box O1A sequence.
XX
KW ss; cytostatic; antidiabetic; foxhead box O1A inhibitor;
KW forkhead box O1A; hyperproliferative disorder; cancer; rhabdomyosarcoma;
KW diabetes; H-ras gene; antisense; gene expression.
XX
OS Mus musculus.
XX
FN WO2004031350-A2.
XX
PD 15-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030352.
XX
PR 26-SEP-2002; 2002US-00260203.
XX
PA (AMGE-) AMGEN INC.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;
XX WPI; 2004-330164/30.
XX
PT New compounds, particularly antisense oligonucleotides, targeted to a
PT nucleic acid molecule encoding forkhead box O1A, useful for treating
PT cancer, or type 2 diabetes.
XX
PS Example 15; SEQ ID NO 11; 146pp; English.
XX
CC The invention relates to a compound 8-80 nucleobases in length targeted
CC to a nucleic acid molecule encoding forkhead box O1A, where the compound
CC is at least 70% complementary to a nucleic acid molecule encoding
CC forkhead box O1A and modulates expression of forkhead box O1A by at least
CC 10%. The compound is useful for treating an animal having a disease or
CC condition associated with forkhead box O1A, e.g. a hyperproliferative
CC disorder (cancer, preferably rhabdomyosarcoma), or type 2 diabetes. This
CC sequence corresponds to the mouse forkhead box O1A to which the
CC oligonucleotides of the invention are targeted.
XX
SQ Sequence 4945 BP; 1148 A; 1303 C; 1209 G; 1285 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 4945;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20
Db 2339 GCTTTGGTTGGCAACACAT 2320

RESULT 14
ADN40657/c
ID ADN40657 standard; DNA; 4945 BP.
XX
AC ADN40657;
XX
DT 12-AUG-2004 (first entry)
XX
DE Mouse forkhead box O1A DNA.
XX
KW Mouse; forkhead box O1A; gene; ds; antisense oligonucleotide;
KW phosphorothioate linkage; 2'-O-methoxyethyl sugar moiety;
KW 5-methylcytosine; hyperproliferative disorder; cancer; rhabdomyosarcoma;
KW type 2 diabetes; cytostatic; antidiabetic.

Mus musculus.
Key Location/Qualifiers
CDS 429..2387
/*tag= a
/product= "Mouse forkhead box O1A"

US2004097459-A1.
20-MAY-2004.
25-SEP-2003; 2003US-00671074.
26-SEP-2002; 2002US-00260203.
(DOBIE/) DOBIE K W.
(BHANOT/) BHANOT S.
(VENIANT-ELLISON M.
(LINDBERG R A.
(SHUTTER J R.
Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;
WPI; 2004-389194/36.
P-PSDB; ADN40824.
GENBANK; AJ252157.
New compounds, particularly antisense oligonucleotides, targeted to a
nucleic acid molecule encoding forkhead box O1A, useful for treating
cancer, or type 2 diabetes.
Example 15; SEQ ID NO 11; 80pp; English.
The invention relates to a compound targeted to a nucleic acid molecule
encoding the human forkhead box O1A polypeptide. The compound is an
antisense oligonucleotide that specifically hybridizes with the nucleic
acid and inhibits expression of the polypeptide. The antisense
oligonucleotide comprises at least one modified internucleoside linkage
i.e. a phosphorothioate linkage, at least one modified sugar moiety,
preferably a 2'-O-methoxyethyl sugar moiety, or at least one modified
nucleobase comprising a 5-methylcytosine. The antisense compounds are
useful for modulating the expression of the human forkhead box O1A
polypeptide and in preparation of a composition for treating
hyperproliferative disorders, e.g. cancer, preferably rhabdomyosarcoma,
and type 2 diabetes. This sequence represents DNA encoding the mouse
forkhead O1A polypeptide of the invention.
Sequence 4945 BP; 1148 A; 1303 C; 1209 G; 1285 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 4945;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20
Db 2339 GCTTTGGTTGGCAACACAT 2320

RESULT 15
ACF87489/c
ID ACF87489 standard; DNA; 5722 BP.
XX
AC ACF87489;
XX
DT 02-JUN-2005 (first entry)
XX
DE Human SIRS/sepsis diagnostic marker DNA fragment 6349.
XX
KW Systemic inflammatory response syndrome; SIRS; antibacterial;
KW immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX
OS Homo sapiens.
```

```
XX WO2004087949-A2.
XX
XX PD
XX PF
XX PP
XX 14-OCT-2004.
XX
XX 31-MAR-2004; 2004WO-EP003419.
XX
XX 02-APR-2003; 2003DE-01015031.
XX
XX 08-AUG-2003; 2003DE-01036511.
XX
XX 02-SEP-2003; 2003DE-01040395.
XX
XX (SIRS-) SIRS LAB GMBH.
XX
XX Russwurm S, Reinhart K, Saluz H, Straube E, Zipfel PF, Deigner H;
XX WPI; 2004-748070/73.
XX
XX In vitro detection of systemic inflammatory response syndrome and related
XX conditions, for e.g. monitoring progression, comprises detecting abnormal
XX expression of disease-related genes.
XX
XX Disclosure; Page; 75pp; German.
XX
XX The invention relates to a novel method for in vitro detection of
XX systemic inflammatory response syndrome (SIRS). The method comprises
XX detecting abnormal expression of disease-related genes, or their
XX associated peptides. The method of the invention demonstrates
XX antibacterial, immunosuppressive and antiinflammatory applications and
XX may be used for early differential diagnosis, monitoring progression,
XX assessing risk, assessing the likely response to treatment and for post
XX mortem diagnosis of systemic inflammatory response syndrome, sepsis and
XX sepsis-like conditions. The recombinant or synthetic nucleic acid
XX sequences of the invention, or derived proteins or peptides, may be
XX useful as calibrants in assays for the specified diseases, for evaluating
XX activity or toxicity in screening for active agents and/or for
XX preparation of agents for treatment or prevention of the specified
XX diseases. The current sequence is that of a human SIRS/sepsis diagnostic
XX marker DNA fragment of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at ftp.wipo.int/pub/published
XX pct sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
XX disclosed within the specification, however, these have not been taken
XX into account during indexing due to inconsistencies in application and
XX format
XX
XX Sequence 5722 BP; 1454 A; 1366 C; 1290 G; 1612 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 13; Length 5722;
XX Best Local Similarity 100.0%; Pred. No. 0.092;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCTTTGGTTGGCAACACAT 20
XX ||||||||||||||||||
XX Db 2305 GCTTTGGTTGGCAACACAT 2286
XX
XX RESULT 16
XX ACF87572/c
XX ID ACF87572 standard; DNA; 5722 BP.
XX
XX AC ACF87572;
XX
XX DT 02-JUN-2005 (first entry)
XX
XX DE Human SIRS/sepsis diagnostic marker DNA fragment 6432.
XX
XX KW Systemic inflammatory response syndrome; SIRS; antibacterial;
XX immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO2004087949-A2.
XX
XX
```

```
PD 14-OCT-2004.
XX
XX 31-MAR-2004; 2004WO-EP003419.
XX
XX 02-APR-2003; 2003DE-01015031.
XX
XX 08-AUG-2003; 2003DE-01036511.
XX
XX 02-SEP-2003; 2003DE-01040395.
XX
XX (SIRS-) SIRS LAB GMBH.
XX
XX Russwurm S, Reinhart K, Saluz H, Straube E, Zipfel PF, Deigner H;
XX WPI; 2004-748070/73.
XX
XX In vitro detection of systemic inflammatory response syndrome and related
XX conditions, for e.g. monitoring progression, comprises detecting abnormal
XX expression of disease-related genes.
XX
XX Disclosure; Page; 75pp; German.
XX
XX The invention relates to a novel method for in vitro detection of
XX systemic inflammatory response syndrome (SIRS). The method comprises
XX detecting abnormal expression of disease-related genes, or their
XX associated peptides. The method of the invention demonstrates
XX antibacterial, immunosuppressive and antiinflammatory applications and
XX may be used for early differential diagnosis, monitoring progression,
XX assessing risk, assessing the likely response to treatment and for post
XX mortem diagnosis of systemic inflammatory response syndrome, sepsis and
XX sepsis-like conditions. The recombinant or synthetic nucleic acid
XX sequences of the invention, or derived proteins or peptides, may be
XX useful as calibrants in assays for the specified diseases, for evaluating
XX activity or toxicity in screening for active agents and/or for
XX preparation of agents for treatment or prevention of the specified
XX diseases. The current sequence is that of a human SIRS/sepsis diagnostic
XX marker DNA fragment of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at ftp.wipo.int/pub/published
XX pct sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
XX disclosed within the specification, however, these have not been taken
XX into account during indexing due to inconsistencies in application and
XX format
XX
XX Sequence 5722 BP; 1454 A; 1366 C; 1290 G; 1612 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 13; Length 5722;
XX Best Local Similarity 100.0%; Pred. No. 0.092;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCTTTGGTTGGCAACACAT 20
XX ||||||||||||||||||
XX Db 2305 GCTTTGGTTGGCAACACAT 2286
XX
XX RESULT 17
XX ABV94143/c
XX ID ABV94143 standard; cDNA; 5723 BP.
XX
XX AC ABV94143;
XX
XX DT 08-JAN-2003 (first entry)
XX
XX DE Breast carcinoma related nucleotide sequence SEQ ID NO:134.
XX
XX KW Human; breast carcinoma; cancer; tumour; cytostatic; anti-tumour; gene;
XX ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200246467-A2.
XX
XX PD 13-JUN-2002.
XX
XX 07-DEC-2001; 2001WO-IB002811.
XX
```

XX 08-DEC-2000; 2000US-0254090P.  
PR 07-DEC-2001; 2001US-00007926.  
XX (IPSO-) IPSOGEN.  
XX Bertucci F, Houlgatte R, Birnbaum D, Nguyen C, Viens P, Fert V;  
FI WPI; 2002-619023/66.  
XX Novel polynucleotide library useful in molecular characterization of a  
PT carcinoma, comprising a pool of polynucleotide sequences or its  
PT subsequence which are either underexpressed or overexpressed in tumor  
PT cells.  
XX Claim 1; Page 198-200; 401pp; English.  
XX The present invention describes a polynucleotide library (I) useful in  
CC the molecular characterisation of a carcinoma, comprising a pool of  
CC polynucleotides or its subsequences which are either underexpressed or  
CC overexpressed in tumour cells, and correspond to any of the  
CC polynucleotide sequences chosen from the 468 sequences given in ABV94010  
CC to ABV94477. Also described: (I) a polynucleotide array (II) useful for  
CC the prognosis or diagnostic of tumour, comprising (I); and (2) detecting  
CC (MI) differentially expressed polynucleotide sequences which are  
CC correlated with a cancer, involves obtaining a polynucleotide sample from  
CC a patient, and reacting the polynucleotide sample obtained with a probe  
CC immobilised on a solid support, where the probe comprises any combination  
CC of the polynucleotide sequences of (I) or its expression products encoded  
CC by polynucleotide sequences of (I), and detecting the reaction product.  
CC (I) have cytostatic activities and can be used as anti-tumour agents. (I)  
CC is useful in molecular characterisation of a carcinoma. (I) and (II) are  
CC useful for the prognosis or diagnostic of tumour, in differentiating a  
CC normal cell from a cancer cell, detecting a hormone sensitive tumour  
CC cell, differentiating a tumour with lymph nodes from a tumour without  
CC lymph nodes, differentiating antitumor agents from tumours from  
CC antitumor agents, differentiating antitumor agents from tumours from  
CC primary breast tumours. (I) is useful for large-scale molecular  
CC characterisation of breast cancer that help in prediction, prognosis and  
CC cancer treatment, and for detecting differentially expressed genes that  
CC correlated with a cancer  
XX  
SQ Sequence 5723 BP; 1454 A; 1366 C; 1291 G; 1612 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 6; Length 5723;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCTTTGGTTGGGCAACACAT 20  
Db 2305 GCTTTGGTTGGGCAACACAT 2286  
RESULT 18  
ADN06031/c  
ID ADN06031 standard; cDNA; 5723 BP.  
XX  
AC ADN06031;  
XX  
XX 01-JUL-2004 (first entry)  
XX Antipsoriatic cDNA sequence #1254.  
XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.  
XX Homo sapiens.  
OS WO2004028479-A2.  
XX  
XX 08-APR-2004.  
PD  
XX 25-SEP-2003; 2003WO-US030907.  
XX

PR 25-SEP-2002; 2002US-0414006P.  
XX (GETH ) GENENTECH INC.  
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;  
FI Wu TD;  
XX WPI; 2004-305105/28.  
DR P-PSDB; ADN06032.  
XX New PRO nucleic acid or polypeptide, useful for preparing a  
PT pharmaceutical composition for diagnosing or treating psoriasis in a  
PT mammal.  
XX Claim 1; SEQ ID NO 2426; 3069pp; English.  
XX The invention relates to novel polynucleotide and polypeptides for  
CC treating psoriasis or a sequence having at least 80% identity to the  
CC above sequences. The nucleic acid is useful for preparing a composition  
CC for diagnosing or treating psoriasis in a mammal. This sequence  
CC corresponds to one of the polynucleotides of the invention.  
XX  
SQ Sequence 5723 BP; 1454 A; 1366 C; 1291 G; 1612 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 12; Length 5723;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCTTTGGTTGGGCAACACAT 20  
Db 2305 GCTTTGGTTGGGCAACACAT 2286  
RESULT 19  
ADN97714/c  
ID ADN97714 standard; DNA; 5723 BP.  
XX  
AC ADN97714;  
XX  
DT 01-JUL-2004 (first entry)  
XX Human foxhead box O1a sequence.  
DE  
XX ss; cytostatic; antidiabetic; foxhead box O1a inhibitor;  
XX forkhead box O1a; hyperproliferative disorder; cancer; rhabdomyosarcoma;  
XX diabetes; H-ras gene; antisense; gene expression.  
XX Homo sapiens.  
OS WO2004031350-A2.  
XX  
XX 15-APR-2004.  
PD  
XX 25-SEP-2003; 2003WO-US030352.  
PF  
XX 26-SEP-2002; 2002US-00260203.  
XX (AMGE-) AMGEN INC.  
PA (ISIS-) ISIS PHARM INC.  
XX  
XX Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;  
FI WPI; 2004-330164/30.  
XX  
XX New compounds, particularly antisense oligonucleotides, targeted to a  
PT nucleic acid molecule encoding forkhead box O1a, useful for treating  
PT cancer, or type 2 diabetes.  
XX  
XX Example 15; SEQ ID NO 4; 146pp; English.  
PS  
XX The invention relates to a compound 8-80 nucleobases in length targeted  
CC to a nucleic acid molecule encoding forkhead box O1a, where the compound  
CC is at least 70% complementary to a nucleic acid molecule encoding

CC forkhead box O1A and modulates expression of forkhead box O1A by at least  
 CC 10%. The compound is useful for treating an animal having a disease or  
 CC condition associated with forkhead box O1A, e.g. a hyperproliferative  
 CC disorder (cancer, preferably rhabdomyosarcoma), or type 2 diabetes. This  
 CC sequence corresponds to the human forkhead box O1A to which the  
 CC oligonucleotides of the invention are targeted.  
 XX

SQ Sequence 5723 BP; 1454 A; 1366 C; 1291 G; 1612 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 5723;  
 Best Local Similarity 100.0%; Pred. No. 0.092;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAACACAT 20  
 |||||  
 Db 2305 GCTTTGGTTGGGCAACACAT 2286

## RESULT 20

ADN40650/C

ID ADN40650 standard; DNA; 5723 BP.

XX AC ADN40650;

XX DT 12-AUG-2004 (first entry)

XX DE Human forkhead box O1A DNA.

XX KW Human; forkhead box O1A; gene; ds; antisense oligonucleotide;

XX KW phosphorothioate linkage; 2'-O-methoxyethyl sugar moiety;

XX KW 5-methylcytosine; hyperproliferative disorder; cancer; rhabdomyosarcoma;

XX KW type 2 diabetes; cytostatic; antidiabetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT 386..2353

XX FT /\*tag= a

XX FT /product= "Human forkhead box O1A"

XX PN US2004097459-A1.

XX PD 20-MAY-2004.

XX PF 25-SEP-2003; 2003US-00671074.

XX PR 26-SEP-2002; 2002US-00260203.

XX PA (DORI/) DOBIE K W.

XX PA (BHANI/) BHANOT S.

XX PA (VENI/) VENIANT-ELLISON M.

XX PA (LIND/) LINDBERG R A.

XX PA (SHUT/) SHUTTER J R.

XX PI Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;

XX WPI; 2004-389194/36.

XX DR P-PSDB; ADN40823.

XX DR GENBANK; NM\_002015.

XX New compounds, particularly antisense oligonucleotides, targeted to a  
 PT nucleic acid molecule encoding forkhead box O1A, useful for treating  
 PT cancer, or type 2 diabetes.

PS Example 15; SEQ ID NO 4; 80pp; English.

XX The invention relates to a compound targeted to a nucleic acid molecule  
 CC encoding the human forkhead box O1A polypeptide. The compound is an  
 CC antisense oligonucleotide that specifically hybridises with the nucleic  
 CC acid and inhibits expression of the polypeptide. The antisense  
 CC oligonucleotide comprises at least one modified internucleoside linkage  
 CC i.e. a phosphorothioate linkage, at least one modified sugar moiety,  
 CC preferably a 2'-O-methoxyethyl sugar moiety, or at least one modified

CC nucleobase comprising a 5-methylcytosine. The antisense compounds are  
 CC useful for modulating the expression of the human forkhead box O1A  
 CC polypeptide and in preparation of a composition for treating  
 CC hyperproliferative disorders, e.g. cancer, preferably rhabdomyosarcoma,  
 CC and type 2 diabetes. This sequence represents DNA encoding the human  
 CC forkhead O1A polypeptide of the invention.  
 XX

SQ Sequence 5723 BP; 1454 A; 1366 C; 1291 G; 1612 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 5723;  
 Best Local Similarity 100.0%; Pred. No. 0.092;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAACACAT 20  
 |||||  
 Db 2305 GCTTTGGTTGGGCAACACAT 2286

## RESULT 21

ACN39628/C

ID ACN39628 standard; cDNA; 5723 BP.

XX AC ACN39628;

XX DT 18-NOV-2004 (first entry)

XX DE Tumour-associated antigenic target (TAT) cDNA DNA270254, SEQ ID NO:3896.

XX KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;

XX KW tumour; diagnosis; cell proliferative disorder; breast cancer;

XX KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;

XX KW central nervous system cancer; bladder cancer; pancreatic cancer;

XX KW cervical cancer; melanoma; leukaemia; hybridisation probe;

XX KW chromosome identification; chromosome mapping; gene mapping;

XX KW gene therapy; cytostatic; gene; ss.

XX OS Homo sapiens.

XX PN WO2004030615-A2.

XX PD 15-APR-2004.

XX PF 29-SEP-2003; 2003WO-US028547.

XX PR 02-OCT-2002; 2002US-0414971P.

XX PA (GETH ) GENENTECH INC.

XX PI Wu TD, Zhang Z, Zhou Y;

XX WPI; 2004-347921/32.

XX DR P-PSDB; ABM81512.

XX New tumor-associated antigenic target polypeptides and nucleic acids,  
 PT useful in preparing a medicament for treating or detecting a  
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
 PT prostate cancer or tumor.

PS Claim 1; SEQ ID NO 3896; 7273pp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)  
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
 CC overexpressed in cancer tissues compared to normal tissues, and may thus  
 CC serve as effective targets for the diagnosis and treatment of cancer in  
 CC mammals. The invention also relates to nucleic acid and polypeptide  
 CC sequences at least 80% identical to the TAT nucleic acids and  
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
 CC TAT polypeptide; and methods and compositions for the treatment or  
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
 CC antibodies, antagonists, binding molecules and compositions are useful  
 CC for diagnosing or treating a cell proliferative disorder associated with

CC increased TAT expression, particularly cancers such as breast cancer, colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder cancer, pancreatic cancer, cervical cancer, cancers of the central nervous system, melanoma and leukemia. TAT nucleic acids may further be used as hybridisation probes, in chromosome and gene mapping, in chromosome identification and in gene therapy. The present sequence CC represents a TAT nucleic acid of the invention

XX SQ Sequence 5723 BP; 1454 A; 1366 C; 1291 G; 1612 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 13; Length 5723;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20  
|||||  
DB 2305 GCTTTGGTTGGCAACACAT 2286

RESULT 22  
ADY14881/c  
ID ADY14881 standard; DNA; 5723 BP.

XX AC ADY14881;

XX DT 05-MAY-2005 (first entry)

XX DE DNA encoding a PRO polypeptide, SEQ ID NO 687.

XX KW Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;  
XX KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;  
XX KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;  
XX KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antisthmatic;  
XX KW Antiallergic; ds; gene; diagnosis.

XX OS Homo sapiens.

XX PN WO2005016962-A2.

XX PD 24-FEB-2005.

XX PF 11-AUG-2004; 2004WO-US026249.

XX PR 11-AUG-2003; 2003US-0493546P.

XX PA (GETH ) GENENTECH INC.

XX PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;

XX DR WPI; 2005-182330/19.

XX PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and treating an immune related disorder, e.g. systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.

XX PS Claim 1; SEQ ID NO 687; 158pp; English.

XX CC The invention relates to an isolated nucleic acid encoding a PRO polypeptide. The polypeptide, agonist or an antagonist, antibody, composition, and method are useful for diagnosing and treating an immune related disorder, e.g. systemic lupus erythematosus, rheumatoid arthritis. The present sequence represents a DNA encoding a PRO polypeptide.

XX SQ Sequence 5723 BP; 1454 A; 1366 C; 1291 G; 1612 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 14; Length 5723;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20  
|||||  
DB 2305 GCTTTGGTTGGCAACACAT 2286

RESULT 23

ADF81591/c

ID ADF81591 standard; DNA; 5769 BP.

XX AC ADF81591;

XX DT 26-FEB-2004 (first entry)

XX DE Leukaemia-related DNA sequence #2147.

XX KW Cytostatic; Gene therapy; leukaemia; ss.

XX OS Unidentified.

XX PN WO2003039443-A2.

XX PD 15-MAY-2003.

XX PF 04-NOV-2002; 2002WO-EP012303.

XX PR 05-NOV-2001; 2001EP-00126244.

XX PR 30-APR-2002; 2002EP-00009758.

XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX PA (UYLU-) UNIV LUDWIG MAXIMILIANS.

XX PA (HAPE/) HAFERLACH T.

XX PA (SCHO/) SCHOCH C.

XX PA (KERN/) KERN W.

XX PI Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;  
XX PI Ellis R, Brors B, Mergenthaler S;

XX DR WPI; 2003-505037/47.

XX PT Determining the subtype of leukemia cells and whether a patient sample contains leukemia cells or other cells, useful for treating leukemia, comprises determining the expression profile of a group of markers in a patient sample.

XX PS Disclosure; SEQ ID NO 2147; 2938pp; English.

XX CC The present invention relates to a method (M1) for determining the subtype of leukaemia cells and whether a patient sample contains leukaemia cells. The method comprises determining the expression profile of a group of markers in a patient sample. The method is useful for determining the presence of leukaemia cells, its types or subtypes, and for the preparation of a medicament for treating leukaemia.

XX SQ Sequence 5769 BP; 1485 A; 1367 C; 1291 G; 1613 T; 0 U; 13 Other;

Query Match 100.0%; Score 20; DB 10; Length 5769;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20  
|||||  
DB 2305 GCTTTGGTTGGCAACACAT 2286

RESULT 24

AAS85053/c

ID AAS85053 standard; cDNA; 5833 BP.

XX AC AAS85053;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #20857.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.



```

XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSB-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG20866.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 20857; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (II) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 5833 BP; 1482 A; 1383 C; 1331 G; 1637 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 5; Length 5833;
XX Best Local Similarity 100.0%; Pred. No. 0.092;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCTTTGGTTGGCACACAT 20
XX
XX Db 2357 GCTTTGGTTGGCACACAT 2338
XX
XX RESULT 25
XX ADN97750
XX ID ADN97750 standard; DNA; 20 BP.
XX AC ADN97750;
XX XX
XX DT 01-JUL-2004 (first entry)
XX DE Human foxhead box O1A sequence inhibitory oligo #23.
XX ss; cytosstatic; antidiabetic; foxhead box O1A inhibitor;
XX KW foxhead box O1A; hyperproliferative disorder; cancer; rhabdomyosarcoma;
XX KW diabetes; H-ras gene; antisense; gene expression; primer.
XX
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT misc_difference 1..20
XX FT /*tag= b
XX FT /note= "sugar phosphate internucleotide linkages in the
XX FT backbone are replaced with a phosphorothioate
XX FT internucleotide linkages"
XX FT modified_base 1..20
XX FT /*tag= c
XX FT /mod_base= OTHER
XX FT /note= "all C are 5'-methylcytidines"
XX FT modified_base 1..5
XX FT /*tag= a
XX FT /mod_base= OTHER
XX FT /note= "nucleotides are 2'-methoxyethyl-nucleotides"
XX FT modified_base 16..20
XX FT /*tag= d
XX FT /mod_base= OTHER
XX FT /note= "nucleotides are 2'-methoxyethyl-nucleotides"
XX XX
XX WO2004031350-A2.
XX PN
XX XX
XX PD 15-APR-2004.
XX XX
XX PF 25-SEP-2003; 2003WO-US030352.
XX PR
XX XX 26-SEP-2002; 2002US-00260203.
XX XX (AMGE-) AMGEN INC.
XX PA (ISIS-) ISIS PHARM INC.
XX XX
XX PI Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;
XX WPI; 2004-330164/30.
XX DR
XX XX New compounds, particularly antisense oligonucleotides, targeted to a
XX PT nucleic acid molecule encoding forkhead box O1A, useful for treating
XX PT cancer, or type 2 diabetes.
XX XX
XX PS Claim 14; SEQ ID NO 40; 146pp; English.
XX XX
XX CC The invention relates to a compound 8-80 nucleobases in length targeted
XX CC to a nucleic acid molecule encoding forkhead box O1A, where the compound
XX CC is at least 70% complementary to a nucleic acid molecule encoding
XX CC forkhead box O1A and modulates expression of forkhead box O1A by at least
XX CC 10%. The compound is useful for treating an animal having a disease or
XX CC condition associated with forkhead box O1A, e.g. a hyperproliferative
XX CC disorder (cancer, preferably rhabdomyosarcoma), or type 2 diabetes. This
XX CC sequence corresponds to an oligonucleotide targeted to the human foxhead
XX CC box O1A genes in order to inhibit gene expression.
XX SQ Sequence 20 BP; 4 A; 4 C; 5 G; 7 T; 0 U; 0 Other;
XX
XX Query Match 95.0%; Score 19; DB 12; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 0.37;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 CTTTGGTTGGCACACAT 20
XX
XX Db 1 CTTTGGTTGGCACACAT 19
XX
XX RESULT 26
XX ADN97828/C
XX ID ADN97828 standard; DNA; 20 BP.
XX XX
XX AC ADN97828;
XX XX
XX DT 01-JUL-2004 (first entry)
XX DE Human foxhead box O1A gene target sequence #21.
XX

```

KW ss; cytostatic; antidiabetic; foxhead box O1A inhibitor;  
KW forkhead box O1A; hyperproliferative disorder; cancer; rhabdomyosarcoma;  
KW diabetes; H-ras gene; antisense; gene expression; primer.  
OS Homo sapiens.  
XX WO2004031350-A2.  
FN 15-APR-2004.  
XX 25-SEP-2003; 2003WO-US030352.  
XX 26-SEP-2002; 2002US-00260203.  
PR (AMGE-) AMGEN INC.  
XX (ISIS-) ISIS PHARM INC.  
PA  
XX Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;  
FI WPI; 2004-330164/30.  
XX  
XX New compounds, particularly antisense oligonucleotides, targeted to a  
PT nucleic acid molecule encoding forkhead box O1A, useful for treating  
PT cancer, or type 2 diabetes.  
XX  
XX Example 18; SEQ ID NO 118; 146pp; English.  
XX  
XX The invention relates to a compound 8-80 nucleobases in length targeted  
CC to a nucleic acid molecule encoding forkhead box O1A, where the compound  
CC is at least 70% complementary to a nucleic acid molecule encoding  
CC forkhead box O1A and modulates expression of forkhead box O1A by at least  
CC 10%. The compound is useful for treating an animal having a disease or  
CC condition associated with forkhead box O1A, e.g. a hyperproliferative  
CC disorder (cancer, preferably rhabdomyosarcoma), or type 2 diabetes. This  
CC sequence corresponds to a targeted sequence from the human foxhead box  
CC O1A gene.  
XX  
SQ Sequence 20 BP; 7 A; 5 C; 4 G; 4 T; 0 U; 0 Other;  
Query Match 95.0%; Score 19; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.37; Mismatches 0; Indels 0; Gaps 0;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 CTTTGGTTGGCAACACAT 20  
Db 20 CTTTGGTTGGCAACACAT 2  
RESULT 27  
ADN40686  
ID ADN40686 standard; DNA; 20 BP.  
XX  
XX ADN40686;  
XX  
XX 12-AUG-2004 (first entry)  
XX Human forkhead box O1A DNA antisense oligonucleotide #23.  
XX  
XX Human; forkhead box O1A; ss; antisense oligonucleotide;  
KW phosphorothioate linkage; 2'-O-methoxyethyl sugar moiety;  
KW 5-methylcytosine; hyperproliferative disorder; cancer; rhabdomyosarcoma;  
KW type 2 diabetes; cytostatic; antidiabetic.  
XX  
XX Homo sapiens.  
XX  
XX US2004097459-A1.  
FN 20-MAY-2004.  
XX  
XX 25-SEP-2003; 2003US-00671074.  
XX  
XX 26-SEP-2002; 2002US-00260203.

PA (DOBI/) DOBIE K W.  
PA (BHAN/) BHANOT S.  
PA (VENI/) VENIANT-ELLISON M.  
PA (LIND/) LINDBERG R A.  
PA (SHUT/) SHUTTER J R.  
XX  
XX Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;  
FI WPI; 2004-389194/36.  
XX  
XX New compounds, particularly antisense oligonucleotides, targeted to a  
PT nucleic acid molecule encoding forkhead box O1A, useful for treating  
PT cancer, or type 2 diabetes.  
XX  
XX Claim 14; SEQ ID NO 40; 80pp; English.  
XX  
XX The invention relates to a compound targeted to a nucleic acid molecule  
CC encoding the human forkhead box O1A polypeptide. The compound is an  
CC antisense oligonucleotide that specifically hybridizes with the nucleic  
CC acid and inhibits expression of the polypeptide. The antisense  
CC oligonucleotide comprises at least one modified internucleoside linkage  
CC i.e. a phosphorothioate linkage, at least one modified sugar moiety,  
CC preferably a 2'-O-methoxyethyl sugar moiety, or at least one modified  
CC nucleobase comprising a 5-methylcytosine. The antisense compounds are  
CC useful for modulating the expression of the human forkhead box O1A  
CC polypeptide and in preparation of a composition for treating  
CC hyperproliferative disorders, e.g. cancer, preferably rhabdomyosarcoma,  
CC and type 2 diabetes. This sequence represents an antisense  
CC oligonucleotide targeted to DNA encoding the human forkhead O1A  
CC polypeptide of the invention.  
XX  
SQ Sequence 20 BP; 4 A; 4 C; 5 G; 7 T; 0 U; 0 Other;  
Query Match 95.0%; Score 19; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.37; Mismatches 0; Indels 0; Gaps 0;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 CTTTGGTTGGCAACACAT 20  
Db 1 CTTTGGTTGGCAACACAT 19  
RESULT 28  
ADN40764/C  
ID ADN40764 standard; DNA; 20 BP.  
XX  
XX ADN40764;  
XX  
XX 12-AUG-2004 (first entry)  
XX Human forkhead box O1A DNA antisense oligonucleotide target region #21.  
XX  
XX Human; forkhead box O1A; ss; antisense oligonucleotide;  
KW phosphorothioate linkage; 2'-O-methoxyethyl sugar moiety;  
KW 5-methylcytosine; hyperproliferative disorder; cancer; rhabdomyosarcoma;  
KW type 2 diabetes; cytostatic; antidiabetic.  
XX  
XX Homo sapiens.  
XX  
XX US2004097459-A1.  
FN 20-MAY-2004.  
XX  
XX 25-SEP-2003; 2003US-00671074.  
XX  
XX 26-SEP-2002; 2002US-00260203.  
XX  
XX (DOBI/) DOBIE K W.  
XX (BHAN/) BHANOT S.  
XX (VENI/) VENIANT-ELLISON M.  
XX (LIND/) LINDBERG R A.  
XX (SHUT/) SHUTTER J R.

PI Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;  
XX WPI; 2004-389194/36.  
XX  
XX New compounds, particularly antisense oligonucleotides, targeted to a  
PT nucleic acid molecule encoding forkhead box O1A, useful for treating  
PT cancer, or type 2 diabetes.  
XX  
XX Example 18; SEQ ID NO 118; 80pp; English.  
XX  
XX The invention relates to a compound targeted to a nucleic acid molecule  
CC encoding the human forkhead box O1A polypeptide. The compound is an  
CC antisense oligonucleotide that specifically hybridizes with the nucleic  
CC acid and inhibits expression of the polypeptide. The antisense  
CC oligonucleotide comprises at least one modified internucleoside linkage  
CC i.e. a phosphorothioate linkage, at least one modified sugar moiety,  
CC preferably a 2'-O-methoxyethyl sugar moiety, or at least one modified  
CC nucleobase comprising a 5-methylcytosine. The antisense compounds are  
CC useful for modulating the expression of the human forkhead box O1A  
CC polypeptide and in preparation of a composition for treating  
CC hyperproliferative disorders, e.g. cancer, preferably rhabdomyosarcoma,  
CC and type 2 diabetes. This sequence represents a human forkhead O1A DNA  
CC antisense oligonucleotide target region of the invention.  
XX  
SQ Sequence 20 BP; 7 A; 5 C; 4 G; 4 T; 0 U; 0 Other;  
  
Query Match 95.0%; Score 19; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 CTTTGGTGGGCAACACAT 20  
DB 20 CTTTGGTGGGCAACACAT 2  
|||||  
  
RESULT 29  
ID ABZ16434/c  
XX ABZ16434 standard; DNA; 2000 BP.  
XX  
XX ABZ16434;  
XX  
XX 21-JAN-2003 (first entry)  
XX  
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 4239.  
XX  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
XX Arabidopsis thaliana.  
XX  
XX WO200216655-A2.  
XX  
XX 28-FEB-2002.  
XX  
XX 24-AUG-2001; 2001WO-US026685.  
XX  
XX 24-AUG-2000; 2000US-0227866P.  
XX  
XX 26-JAN-2001; 2001US-0264647P.  
XX  
XX 22-JUN-2001; 2001US-0300111P.  
XX  
XX (SCRI ) SCRIPPS RES INST.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
XX Harper JF, Kreps J, Wang X, Zhu T;  
XX  
XX WPI; 2002-304127/34.  
XX  
XX Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
XX  
XX Claim 144; SEQ ID NO 4239; 577pp + Sequence Listing; English.  
XX  
XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid

CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: the sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
SQ Sequence 2000 BP; 657 A; 326 C; 263 G; 754 T; 0 U; 0 Other;  
  
Query Match 85.0%; Score 17; DB 6; Length 2000;  
Best Local Similarity 100.0%; Pred. No. 4.7;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GCTTTGGTGGGCAACA 17  
DB 327 GCTTTGGTGGGCAACA 311  
|||||  
  
RESULT 30  
ID ADN97751 standard; DNA; 20 BP.  
XX ADN97751  
XX  
XX ADN97751;  
XX  
XX 01-JUL-2004 (first entry)  
XX  
XX Human forkhead box O1A sequence inhibitory oligo #24.  
XX  
XX ss; cytostatic; antidiabetic; forkhead box O1A inhibitor;  
XX forkhead box O1A; hyperproliferative disorder; cancer; rhabdomyosarcoma;  
XX diabetes; H-ras gene; antisense; gene expression; primer.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH misc\_difference 1..20  
FT /\*tag= b  
FT /note= "sugar phosphate internucleotide linkages in the  
FT backbone are replaced with a phosphorothioate  
FT internucleotide linkages"  
FT modified\_base 1..20  
FT /\*tag= c  
FT /mod\_base= OTHER  
FT /note= "all C are 5'-methylcytidines"  
FT modified\_base 1..5  
FT /mod\_base= a  
FT /note= "nucleotides are 2'-methoxyethyl-nucleotides"  
FT modified\_base 16..20  
FT /\*tag= d  
FT /mod\_base= OTHER  
FT /note= "nucleotides are 2'-methoxyethyl-nucleotides"  
XX  
XX WO2004031350-A2.  
XX  
XX 15-APR-2004.  
XX  
XX 25-SEP-2003; 2003WO-US030352.  
XX  
XX 26-SEP-2002; 2002US-00260203.  
XX  
XX (AMGE-) AMGEN INC.  
XX (ISIS-) ISIS PHARM INC.  
XX  
XX Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;  
XX WPI; 2004-330164/30.  
XX  
XX New compounds, particularly antisense oligonucleotides, targeted to a

PT nucleic acid molecule encoding forkhead box O1A, useful for treating  
PT cancer, or type 2 diabetes.  
PS Claim 14; SEQ ID NO 41; 146pp; English.  
XX  
XX The invention relates to a compound 8-80 nucleobases in length targeted  
CC to a nucleic acid molecule encoding forkhead box O1A, where the compound  
CC is at least 70% complementary to a nucleic acid molecule encoding  
CC forkhead box O1A and modulates expression of forkhead box O1A by at least  
CC 10%. The compound is useful for treating an animal having a disease or  
CC condition associated with forkhead box O1A, e.g. a hyperproliferative  
CC disorder (cancer, preferably rhabdomyosarcoma), or type 2 diabetes. This  
CC sequence corresponds to an oligonucleotide targeted to the human forkhead  
CC box O1A genes in order to inhibit gene expression.  
XX  
SQ Sequence 20 BP; 4 A; 3 C; 8 G; 5 T; 0 U; 0 Other;  
  
Query Match 80.0%; Score 16; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GCTTTGGTTGGGCAAC 16  
Db 5 GCTTTGGTTGGGCAAC 20  
  
RESULT 31  
ADN97829/c  
ID ADN97829 standard; DNA; 20 BP.  
XX  
XX AC ADN97829;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Human forkhead box O1A gene target sequence #22.  
XX  
XX ss; cytostatic; antidiabetic; forkhead box O1A inhibitor;  
KW forkhead box O1A; hyperproliferative disorder; cancer; rhabdomyosarcoma;  
KW diabetes; H-ras gene; antisense; gene expression; primer.  
XX  
XX Homo sapiens.  
OS  
XX WO2004031350-A2.  
FN  
XX  
XX 15-APR-2004.  
PD  
XX  
XX 25-SEP-2003; 2003WO-US030352.  
PF  
XX  
XX 26-SEP-2002; 2002US-00260203.  
PR  
XX  
XX (AMGE-) AMGEN INC.  
PA  
XX (ISIS-) ISIS PHARM INC.  
PA  
XX  
PI Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;  
XX  
XX WPI; 2004-330164/30.  
DR  
XX  
XX New compounds, particularly antisense oligonucleotides, targeted to a  
PT nucleic acid molecule encoding forkhead box O1A, useful for treating  
PT cancer, or type 2 diabetes.  
XX  
XX Example 18; SEQ ID NO 119; 146pp; English.  
PS  
XX  
XX The invention relates to a compound 8-80 nucleobases in length targeted  
CC to a nucleic acid molecule encoding forkhead box O1A, where the compound  
CC is at least 70% complementary to a nucleic acid molecule encoding  
CC forkhead box O1A and modulates expression of forkhead box O1A by at least  
CC 10%. The compound is useful for treating an animal having a disease or  
CC condition associated with forkhead box O1A, e.g. a hyperproliferative  
CC disorder (cancer, preferably rhabdomyosarcoma), or type 2 diabetes. This  
CC sequence corresponds to a targeted sequence from the human forkhead box  
CC O1A gene.  
XX

SQ Sequence 20 BP; 5 A; 8 C; 3 G; 4 T; 0 U; 0 Other;  
  
Query Match 80.0%; Score 16; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GCTTTGGTTGGGCAAC 16  
Db 16 GCTTTGGTTGGGCAAC 1  
  
RESULT 32  
ADN40687  
ID ADN40687 standard; DNA; 20 BP.  
XX  
XX AC ADN40687;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human forkhead box O1A DNA antisense oligonucleotide #24.  
XX  
XX Human; forkhead box O1A; ss; antisense oligonucleotide;  
KW phosphorothioate linkage; 2'-O-methoxyethyl sugar moiety;  
KW 5-methylcytosine; hyperproliferative disorder; cancer; rhabdomyosarcoma;  
KW type 2 diabetes; cytostatic; antidiabetic.  
XX  
XX Homo sapiens.  
OS  
XX US2004097459-A1.  
FN  
XX 20-MAY-2004.  
PD  
XX  
XX 25-SEP-2003; 2003US-00671074.  
PF  
XX  
XX 26-SEP-2002; 2002US-00260203.  
PR  
XX  
XX (DOBI/) DOBIE K W.  
PA (BHAN/) BHANOT S.  
PA (VENI/) VENIANT-ELLISON M.  
PA (LIND/) LINDBERG R A.  
PA (SHUT/) SHUTTER J R.  
XX  
PI Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;  
XX  
XX WPI; 2004-389194/36.  
DR  
XX  
XX New compounds, particularly antisense oligonucleotides, targeted to a  
PT nucleic acid molecule encoding forkhead box O1A, useful for treating  
PT cancer, or type 2 diabetes.  
XX  
XX Claim 14; SEQ ID NO 41; 80pp; English.  
PS  
XX  
XX The invention relates to a compound targeted to a nucleic acid molecule  
CC encoding the human forkhead box O1A polypeptide. The compound is an  
CC antisense oligonucleotide that specifically hybridizes with the nucleic  
CC acid and inhibits expression of the polypeptide. The antisense  
CC oligonucleotide comprises at least one modified internucleoside linkage,  
CC i.e. a phosphorothioate linkage, at least one modified sugar moiety,  
CC preferably a 2'-O-methoxyethyl sugar moiety, or at least one modified  
CC nucleobase comprising a 5-methylcytosine. The antisense compounds are  
CC useful for modulating the expression of the human forkhead box O1A  
CC polypeptide and in preparation of a composition for treating  
CC hyperproliferative disorders, e.g. cancer, preferably rhabdomyosarcoma,  
CC and type 2 diabetes. This sequence represents an antisense  
CC oligonucleotide targeted to DNA encoding the human forkhead O1A  
CC polypeptide of the invention.  
XX  
SQ Sequence 20 BP; 4 A; 3 C; 8 G; 5 T; 0 U; 0 Other;  
  
Query Match 80.0%; Score 16; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GCTTTGGTTGGCAAC 16
Db 5 GCTTTGGTTGGCAAC 20

RESULT 33
ADN40765/c
ID ADN40765 standard; DNA; 20 BP.
XX
AC ADN40765;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human forkhead box O1A DNA antisense oligonucleotide target region #22.
XX
KW Human; forkhead box O1A; ss; antisense oligonucleotide;
KW phosphorothioate linkage; 2'-O-methoxyethyl sugar moiety;
KW 5-methylcytosine; hyperproliferative disorder; cancer; rhabdomyosarcoma;
KW type 2 diabetes; cytostatic; antidiabetic.
XX
OS Homo sapiens.
XX
PN US2004097459-A1.
XX
PD 20-MAY-2004.
XX
PF 25-SEP-2003; 2003US-00671074.
XX
PR 26-SEP-2002; 2002US-00260203.
XX
PA (DOI/) DOBIE K W.
PA (BHAN/) BHANOT S.
PA (VENI/) VENIANT-ELLISON M.
PA (LIND/) LINDBERG R A.
PA (SHUT/) SHUTTER J R.
XX
PI Dobie KW, Bhanot S, Veniant-Ellison M, Lindberg RA, Shutter JR;
XX WPI; 2004-389194/36.
XX
PT New compounds, particularly antisense oligonucleotides, targeted to a
PT nucleic acid molecule encoding forkhead box O1A, useful for treating
PT cancer, or type 2 diabetes.
XX
PS Example 18; SEQ ID NO 119; 80pp; English.
XX
CC The invention relates to a compound targeted to a nucleic acid molecule
CC encoding the human forkhead box O1A polypeptide. The compound is an
CC antisense oligonucleotide that specifically hybridises with the nucleic
CC acid and inhibits expression of the polypeptide. The antisense
CC oligonucleotide comprises at least one modified internucleoside linkage
CC i.e. a phosphorothioate linkage, at least one modified sugar moiety,
CC preferably a 2'-O-methoxyethyl sugar moiety, or at least one modified
CC nucleobase comprising a 5-methylcytosine. The antisense compounds are
CC useful for modulating the expression of the human forkhead box O1A
CC polypeptide and in preparation of a composition for treating
CC hyperproliferative disorders, e.g. cancer, preferably rhabdomyosarcoma,
CC and type 2 diabetes. This sequence represents a human forkhead O1A DNA
CC antisense oligonucleotide target region of the invention.
XX
SQ Sequence 20 BP; 5 A; 8 C; 3 G; 4 T; 0 U; 0 Other;

Query Match 80.0%; Score 16; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAAC 16
Db 16 GCTTTGGTTGGCAAC 1

RESULT 34
AAL05051
```

```
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249293P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250300P.
PR 01-DEC-2000; 2000US-0250310P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251898P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX Disclosure; SEQ ID NO 7739; 1297pp + Sequence Listing; English.
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX SQ Sequence 4106 BP; 1278 A; 669 C; 703 G; 1456 T; 0 U; 0 Other;
Query Match 80.0%; Score 16; DB 4; Length 4106;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 TTGGTTGGGCACACA 19
Db 1936 TTGGTTGGGCACACA 1951
RESULT 35
ABL97944
ID ABL97944 standard; DNA; 4106 BP.
XX AC ABL97944;
XX 21-JUN-2002 (first entry)
DT Human testicular antigen encoding DNA fragment SEQ ID NO: 2596.
DE Human; testicular antigen; testes; cancer; metastasis; immune disorder;
XX reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX Homo sapiens.
OS WO200155317-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001329.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
```

PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 26-JUL-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-022547P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-023397P.  
 PR 14-SEP-2000; 2000US-023398P.  
 PR 14-SEP-2000; 2000US-023399P.  
 PR 14-SEP-2000; 2000US-023399P.  
 PR 14-SEP-2000; 2000US-023400P.  
 PR 14-SEP-2000; 2000US-023401P.  
 PR 14-SEP-2000; 2000US-023403P.  
 PR 14-SEP-2000; 2000US-023406P.  
 PR 14-SEP-2000; 2000US-023406P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241828P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-483232/52.  
 DR  
 XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
 PT for preventing, diagnosing and/or treating testicular cancer.  
 XX  
 PS Disclosure; SEQ ID NO 2596; 766pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of 973  
 CC human testicular antigens, and fragments of their genomic sequences. The  
 CC sequences can be used in the treatment of cardiovascular, urinary system,  
 CC reproductive system, immune, respiratory, neurological and  
 CC gastrointestinal disorders, infections, and particularly cancer,  
 CC especially testicular cancers. The present sequence is a DNA encoding a  
 CC protein fragment of the invention  
 XX  
 XX Sequence 4106 BP; 1278 A; 669 C; 703 G; 1456 T; 0 U; 0 Other;

```
Query Match      80.0%; Score 16; DB 4; Length 4106;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 TTGGTTGGGCAACACA 19
      |||||
Db      1936 TTGGTTGGGCAACACA 1951

RESULT 36
ADL13962/C
ID ADL13962 standard; DNA; 160198 BP.
XX
XX ADL13962;
AC
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX Osteoarthritis-associated polymorphic nucleotide #494.
DE
XX
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
XX Homo sapiens.
OS
XX
XX WO2003054166-A2.
FN
XX
XX 03-JUL-2003.
PD
XX
XX 19-DEC-2002; 2002WO-US041225.
PF
XX
XX 20-DEC-2001; 2001US-0342603P.
PR
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX Jones KA, Schafer A;
PI
XX
XX WPI; 2003-559141/52.
DR
XX
XX Determining susceptibility of an individual to joint space narrowing,
PT osteophyte development and/or joint pain comprises identifying whether
PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.
XX
XX Disclosure; SEQ ID NO 494; 297pp; English.
PS
XX
XX The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space
CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).
XX
XX SQ Sequence 160198 BP; 48087 A; 32398 C; 32827 G; 46852 T; 0 U; 34 Other;

Query Match      80.0%; Score 16; DB 10; Length 160198;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCTTGGTTGGGCAAC 16
      |||||
Db      157646 GCTTGGTTGGGCAAC 157631

RESULT 37
ADL13904/C
ID ADL13904 standard; DNA; 164772 BP.
XX
XX ADL13904;
AC
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX Osteoarthritis-associated polymorphic nucleotide #436.
DE
XX
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
XX Homo sapiens.
OS
XX
XX WO2003054166-A2.
FN
XX
XX 03-JUL-2003.
PD
XX
XX 19-DEC-2002; 2002WO-US041225.
PF
XX
XX 20-DEC-2001; 2001US-0342603P.
PR
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX Jones KA, Schafer A;
PI
XX
XX WPI; 2003-559141/52.
DR
XX
XX Determining susceptibility of an individual to joint space narrowing,
PT osteophyte development and/or joint pain comprises identifying whether
PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.
XX
XX Disclosure; SEQ ID NO 436; 297pp; English.
PS
XX
XX The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space
CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).
XX
XX SQ Sequence 164772 BP; 50645 A; 32137 C; 31960 G; 50022 T; 0 U; 8 Other;

Query Match      80.0%; Score 16; DB 10; Length 164772;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 TTGGTTGGGCAACACA 19
      |||||
Db      49632 TTGGTTGGGCAACACA 49617

RESULT 38
ADL88066
ID ADL88066 standard; DNA; 406 BP.
XX
XX ADL88066;
AC
XX
XX 20-MAY-2004 (first entry)
DT
XX
XX DNA up-regulated in murine haematopoietic stem cells cells SeqID 4459.
```



XX gene potential; multi-lineage; cell commitment; haematopoietic stem cell;  
 KW HSC; multipotent progenitor; MPP; common lymphoid progenitor; CLP;  
 KW common myeloid progenitor; CMP; bone marrow stem cell; mouse; murine; ds.  
 XX Mus sp.  
 OS WO2003093445-A2.  
 PN 13-NOV-2003.  
 PD  
 XX  
 XX 05-MAY-2003; 2003WO-US014114.  
 PF  
 XX 03-MAY-2002; 2002US-0377383P.  
 PR  
 XX (STOW-) STOWERS INST MEDICAL RES.  
 PA  
 XX Li L;  
 PI WPI; 2004-022656/02.  
 DR  
 XX Classifying an unknown multi-lineage affiliated gene comprises isolating  
 PT expressed nucleic acid sequences from the discrete cell sub-populations.  
 PT  
 XX Claim 7; SEQ ID NO 4459; 123pp; English.  
 PS  
 XX This invention relates to a novel method for predicting gene potential by  
 CC associating nucleic acid sequences of unknown function with particular  
 CC sub-population profiles. Specifically, it refers to classifying an  
 CC unknown multi-lineage affiliated gene by collecting hybridisation data to  
 CC develop a gene expression map, in order to determine the discrete sub-  
 CC population where it is expressed. The present invention describes methods  
 CC for predicting the lineage commitment of genes associated with the self-  
 CC renewing haematopoietic (blood) stem cells (HSCs), as well as the non-  
 CC self renewing multipotent progenitors (MPPs), common lymphoid progenitors  
 CC (CLPs) and common myeloid progenitors (CMPs), which are collectively  
 CC referred to as bone marrow stem cells populations. As such, these methods  
 CC can be used to identify associated multi-lineage affiliated genes and  
 CC hence the underlying molecular mechanisms in physiological haematopoietic  
 CC development. This polynucleotide sequence is DNA associated with a murine  
 CC HSC sub population of cells of the invention.  
 XX  
 SQ Sequence 406 BP; 103 A; 70 C; 94 G; 136 T; 0 U; 3 Other;  
 Query Match 75.0%; Score 15; DB 12; Length 406;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 GGTGGGCAACACAT 20  
 DB |||||  
 267 GGTGGGCAACACAT 281  
 RESULT 39  
 ADL88067  
 ID ADL88067 standard; DNA; 406 BP.  
 XX  
 AC ADL88067;  
 XX  
 XX 20-MAY-2004 (first entry)  
 DT  
 DE DNA up-regulated in murine haematopoietic stem cells SeqID 4460.  
 XX  
 KW gene potential; multi-lineage; cell commitment; haematopoietic stem cell;  
 KW HSC; multipotent progenitor; MPP; common lymphoid progenitor; CLP;  
 KW common myeloid progenitor; CMP; bone marrow stem cell; mouse; murine; ds.  
 XX Mus sp.  
 OS WO2003093445-A2.  
 PN 13-NOV-2003.  
 PD  
 XX

PF 05-MAY-2003; 2003WO-US014114.  
 XX  
 PR 03-MAY-2002; 2002US-0377383P.  
 XX  
 PA (STOW-) STOWERS INST MEDICAL RES.  
 XX  
 XX Li L;  
 PI WPI; 2004-022656/02.  
 DR  
 XX Classifying an unknown multi-lineage affiliated gene comprises isolating  
 PT expressed nucleic acid sequences from the discrete cell sub-populations.  
 PT  
 XX Claim 7; SEQ ID NO 4460; 123pp; English.  
 PS  
 XX This invention relates to a novel method for predicting gene potential by  
 CC associating nucleic acid sequences of unknown function with particular  
 CC sub-population profiles. Specifically, it refers to classifying an  
 CC unknown multi-lineage affiliated gene by collecting hybridisation data to  
 CC develop a gene expression map, in order to determine the discrete sub-  
 CC population where it is expressed. The present invention describes methods  
 CC for predicting the lineage commitment of genes associated with the self-  
 CC renewing haematopoietic (blood) stem cells (HSCs), as well as the non-  
 CC self renewing multipotent progenitors (MPPs), common lymphoid progenitors  
 CC (CLPs) and common myeloid progenitors (CMPs), which are collectively  
 CC referred to as bone marrow stem cells populations. As such, these methods  
 CC can be used to identify associated multi-lineage affiliated genes and  
 CC hence the underlying molecular mechanisms in physiological haematopoietic  
 CC development. This polynucleotide sequence is DNA associated with a murine  
 CC HSC sub population of cells of the invention.  
 XX  
 SQ Sequence 406 BP; 103 A; 70 C; 94 G; 136 T; 0 U; 3 Other;  
 Query Match 75.0%; Score 15; DB 12; Length 406;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 GGTGGGCAACACAT 20  
 DB |||||  
 267 GGTGGGCAACACAT 281  
 RESULT 40  
 ACH45273/c  
 ID ACH45273 standard; cDNA; 498 BP.  
 XX  
 AC ACH45273;  
 XX  
 DT 13-OCT-2003 (first entry)  
 DE  
 DE Human foetal brain cDNA #5998.  
 XX  
 KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
 KW genome mapping; biodiversity; genetic disorder.  
 XX  
 OS Homo sapiens.  
 OS  
 XX US2003073623-A1.  
 PN  
 XX 17-APR-2003.  
 PD  
 XX 30-JUL-2001; 2001US-00918995.  
 PF  
 XX 30-JUL-2001; 2001US-00918995.  
 PR  
 XX (DRMA/) DRMANAC R T.  
 PA (LABA/) LABAT I.  
 PA (STAC/) STACHE-CRAIN B.  
 PA (DICK/) DICKSON M C.  
 PA (JONE/) JONES L W.  
 XX  
 PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
 XX

DR WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful

PT as hybridization probes, as oligomers for PCR, for chromosome and gene

PT mapping, in the recombinant production of protein, or in generating

XX antisense DNA or RNA.

XX Claim 1; SEQ ID NO 32485; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of

CC 38043 cDNA sequences, appearing as ACH1789-ACH50831, whose sequence was

CC determined by the technique of SBH (sequencing by hybridisation). Also

CC included is a purified polypeptide comprising a sequence corresponding to

CC a reading frame of the novel polynucleotide. The nucleic acid sequences

CC are useful in diagnostics as expressed sequence tags (EST) for

CC identifying expressed genes or for physical mapping of the human genome,

CC in forensics, in assessing biodiversity, or in identifying mutations

CC responsible for genetic disorders and other traits. The nucleotide

CC sequences are also useful as hybridisation probes, as oligomers for PCR,

CC for chromosome and gene mapping, in the recombinant production of

CC protein, or in generating antisense DNA or RNA. The purified polypeptide

CC is useful for generating antibodies specific for it. The present sequence

CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data

CC for this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from USPTO at

XX seqdata.uspto.gov/sequence.html?DocID=20030073623

XX SQ Sequence 498 BP; 128 A; 139 C; 136 G; 86 T; 0 U; 9 Other;

Query Match 75.0%; Score 15; DB 9; Length 498;

Best Local Similarity 100.0%; Pred. No. 67;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAA 15

Db 253 GCTTTGGTTGGGCAA 239

RESULT 41

AAK88397

ID AAK88397 standard; cDNA; 574 BP.

XX AC AAK88397;

XX DT 05-NOV-2001 (first entry)

XX DE Human digestive system antigen coding sequence SEQ ID NO: 713.

XX Human; digestive system antigen; gene therapy; cancer; appendicitis;

KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;

KW digestive system disorder; Meckel's diverticulum; ss.

XX OS Homo sapiens.

XX WO200155314-A2.

XX FN 02 -AUG-2001.

XX PD 17-JAN-2001; 2001WO-US001324.

XX PF 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226681P.

PR 23-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 12-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 13-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI; 2001-502630/55.  
DR P-PSDB; AAM92624.  
XX  
XX Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases.  
XX  
XX Claim 1; SEQ ID NO 713; 986pp; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a cDNA encoding a digestive  
XX system antigen of the invention  
XX  
XX Sequence 574 BP; 148 A; 113 C; 137 G; 172 T; 0 U; 4 Other;  
SQ

Query Match 75.0%; Score 15; DB 4; Length 574;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 TTGTGTTGGCAACA 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 539 TTGTGTTGGCAACA 553  
RESULT 42  
AAS39498  
ID AAS39498 standard; cDNA; 574 BP.  
XX AAS39498;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE cDNA encoding novel human colon associated polypeptide #151.  
XX  
KW Human; colon cancer; congenital abnormality; infection; colitis;  
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;  
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;  
KW sigmoid disease; antibacterial; antiviral; antiinflammatory; cytostatic;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200155302-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001240.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 06-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI; 2001-465567/50.  
DR P-PSDB; AAU22618.  
XX  
PT Isolated polypeptide for treating, preventing and/ or prognosing  
PT disorders related to the colon including colon cancers and also for  
PT testing and detection e.g. diagnosis.  
XX  
PS Claim 4; SEQ ID NO 161; 562pp; English.  
XX  
CC The present invention relates to the isolation of novel human colon  
CC associated polypeptides (AAU2248-AAU22701), and the cDNA and genomic  
CC sequences encoding for them. The sequences of the invention are useful in  
CC the diagnosis, treatment, prevention and/or prognosis of disorders of the  
CC colon including colon cancer, congenital abnormalities (e.g. atresia and  
CC stenosis), bacterial and viral infections, inflammatory bowel disease  
CC (IBD), neoplastic cell disorders (e.g. polyps and adenomas, intestinal  
CC inflammatory disorders, colitis, colonic inflammation, diarrhoea and  
CC dysentery, malabsorption syndromes (e.g. lactose intolerance), intestinal  
CC obstruction and sigmoid diseases. The polynucleotide sequences of the  
CC invention can also be used in gene therapy. AAS39348-AAS39581 represent  
CC cDNA sequences encoding for the novel human colon associated polypeptides  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 574 BP; 148 A; 113 C; 137 G; 172 T; 0 U; 4 Other;  
Query Match 75.0%; Score 15; DB 5; Length 574;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 TTTGGTTGGGCAACA 17  
|||||  
Db 539 TTTGGTTGGGCAACA 553  
RESULT 43  
ADB32224  
ID ADB32224 standard; cDNA; 574 BP.  
XX  
AC ADB32224;  
XX  
DT 04-DEC-2003 (first entry)  
XX

DE Human novel colon related polypeptide cDNA SEQ ID NO 161.  
XX gene therapy; ss; gene; cancer; liver disorder; hepatitis;  
KW neural disorder; Alzheimer's disease; human; colon.  
XX Homo sapiens.  
OS  
XX  
PN US2003050231-A1.  
XX 13-MAR-2003.  
PD  
XX  
XX  
XX 17-JAN-2001; 2001US-00764872.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226273P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231245P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240606P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 08-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249219P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.

```
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2003-625420/59.
DR P-PSDB; ADB32458.
DR
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX
XX Claim 3; SEQ ID NO 161; 216pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule encoding a
CC polypeptide. The nucleic acid is useful for preparing a medicament for
CC preventing, treating or ameliorating a medical condition e.g. cancer,
CC liver disorders such as hepatitis or neural disorders such as Alzheimer's
CC disease. The present sequence represents a human cDNA encoding a novel
CC colon related polypeptide. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030050231.
XX
XX Sequence 574 BP; 148 A; 113 C; 137 G; 172 T; 0 U; 4 Other;
SQ
Query Match 75.0%; Score 15; DB 9; Length 574;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 TTTGTTGGGCAACA 17
Db 539 TTTGTTGGGCAACA 553
RESULT 44
AAS75868/C
ID AAS75868 standard; cDNA; 1263 BP.
XX
XX AAS75868;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
DE DNA encoding novel human diagnostic protein #11672.
XX
XX Human; chromosome mapping; Gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG11681.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
PT
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 11672; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activities. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1263 BP; 403 A; 292 C; 361 G; 207 T; 0 U; 0 Other;
SQ
Query Match 75.0%; Score 15; DB 5; Length 1263;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCTTTGGTTGGCAA 15
Db 255 GCTTTGGTTGGCAA 241
RESULT 45
AAS31252
ID AAS31252 standard; DNA; 1342 BP.
XX
XX AAS31252;
AC
XX
XX 01-OCT-1998 (first entry)
DT
XX
DE E. coli J96 pathogenicity island contig #66.
XX
XX PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pHER;
XX PAI V; phev; vaccine; protective immune response; ds.
XX
XX Escherichia coli.
OS
XX
XX WO9822575-A2.
XX
XX 28-MAY-1998.
XX
XX 21-NOV-1997; 97WO-US021347.
XX
XX 22-NOV-1996; 96US-0031626P.
XX
XX 14-OCT-1997; 97US-0061953P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX (UYWI-) UNIV WISCONSIN.
XX
XX Dillon FJ, Choi GH, Welch RA;
XX
XX WPI; 1998-312461/27.
XX
XX New isolated uropathogenic E. coli nucleotide sequences - used to develop
XX products for the detection of pathogenic E. coli and to elicit an immune
XX response to pathogenic E. coli.
XX
```

PS Claim 21; Page 167-168; 250pp; English.

XX This sequence represents a E. coli strain J96 contig containing  
CC pathogenicity island (PAI) sequences, and represents a nucleic acid  
CC molecule of the invention. PAIs are large fragments of DNA which comprise  
CC pathogenicity determinants. The sequences of the invention are taken from  
CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)  
CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at  
CC approximately 94 min (at pheK) on the E. coli chromosome and is  
CC approximately 160 kb in size. Antibodies specific to the proteins encoded  
CC by the PAI open reading frames of the invention can be used in kits to  
CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit  
CC a protective immune response in an animal to the uropathogenic E. coli  
CC strain J96

SQ Sequence 1342 BP; 368 A; 319 C; 266 G; 381 T; 0 U; 8 Other;  
Query Match 75.0%; Score 15; DB 2; Length 1342;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAA 15  
|||||  
Db 614 GCTTTGGTTGGCAA 628  
|||||

RESULT 46  
ADM02861/c  
ID ADM02861 standard; cDNA; 1550 BP.  
XX  
AC ADM02861;  
XX  
XX 20-MAY-2004 (first entry)  
XX Human cDNA of the invention SEQ ID NO:1546.  
XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.  
XX Homo sapiens.  
XX EP1347046-A1.  
XX  
XX 24-SEP-2003.  
XX  
XX 12-APR-2002; 2002EP-00008400.  
XX  
XX 22-MAR-2002; 2002JP-00137785.  
XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
XX WPI; 2003-723558/69.  
DR P-PSDB; ADM05304.  
XX  
XX New polynucleotides and polypeptides are useful in gene therapy, for  
PT developing a diagnostic marker or medicines for regulating their  
PT expression and activity, or as a target of gene therapy.  
XX  
XX Claim 1; SEQ ID NO 1546; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded  
CC polypeptide. A polynucleotide of the invention may have a use in gene  
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
CC as a primer for synthesizing the polynucleotide or as a probe for  
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
CC useful in gene therapy, for developing a diagnostic marker or medicines  
CC for regulating their expression and activity, or as a target of gene  
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
CC are useful as pharmaceutical agents. The present sequence represents a  
CC cDNA sequence of the invention.

XX Sequence 1550 BP; 495 A; 332 C; 423 G; 300 T; 0 U; 0 Other;  
SQ  
Query Match 75.0%; Score 15; DB 11; Length 1550;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAA 15  
|||||  
Db 269 GCTTTGGTTGGCAA 255  
|||||

RESULT 47  
ABQ75898  
ID ABQ75898 standard; cDNA; 1635 BP.  
XX  
AC ABQ75898;  
XX  
XX 17-OCT-2002 (first entry)  
XX Human ubiquitin relative protein 46.64 cDNA.  
XX  
XX Human; ubiquitin relative protein 46.64; tumour; inflammation;  
KW immunological disease; haemopathy; human immunodeficiency virus; HIV;  
KW gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 55..1329  
FT /\*tag= a  
FT /product= "ubiquitin relative protein 46.64"  
XX  
XX CN1339485-A.  
XX  
XX 13-MAR-2002.  
XX  
XX 23-AUG-2000; 2000CN-00119708.  
XX  
XX 23-AUG-2000; 2000CN-00119708.  
XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.  
XX  
XX Mao Y, Xie Y;  
XX  
XX WPI; 2002-464069/50.  
DR P-PSDB; ABQ75898.  
XX  
XX New polypeptide-human ubiquitin protein relative protein 46.64 for  
PT treating malignant tumors, inflammations, immunological diseases,  
PT hemopathy and human immunodeficiency virus infection.  
XX  
XX Claim 6; Page 24-25 (disclosure); 33pp; Chinese.

XX The present invention discloses a new kind of polypeptide, human  
CC ubiquitin relative protein 46.64, polynucleotides for encoding this  
CC polypeptide, and a DNA recombination process to produce the polypeptide.  
CC The present invention also discloses the method of applying the  
CC polypeptide in treating various diseases, such as malignant tumours,  
CC inflammations, immunological diseases, haemopathy and human  
CC immunodeficiency virus (HIV) infection. The current sequence represents  
CC the human ubiquitin relative protein 46.64 cDNA

SQ Sequence 1635 BP; 529 A; 290 C; 344 G; 472 T; 0 U; 0 Other;  
Query Match 75.0%; Score 15; DB 6; Length 1635;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GCTTGGGCAACACAT 20  
|||||  
Db 132 GCTTGGGCAACACAT 146  
|||||

RESULT 48  
ADF03345  
ID ADF03345 standard; DNA; 1701 BP.  
XX  
XX  
AC ADF03345;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
XX Bacterial polynucleotide #3630.  
XX  
XX Proteus mirabilis infection; bacterial infection; antibacterial;  
KW immuno stimulant; gene; ds.  
XX  
XX Proteus mirabilis.  
XX  
XX US6605709-B1.  
PN  
XX  
PD 12-AUG-2003.  
XX  
XX 05-APR-2000; 2000US-00543681.  
PF  
XX 09-APR-1999; 99US-0128706P.  
PR  
XX (GENO-) GENOME THERAPEUTICS CORP.  
PA  
PI Breton GL;  
XX  
XX WPI; 2003-895291/82.  
DR P-PSDB; ADF07517.  
XX  
XX New Proteus mirabilis polypeptides and polynucleotides, useful as  
PT reagents for diagnosis of bacterial disease, as components of  
PT antibacterial vaccines, as targets for antibacterial drugs, or as  
PT biocontrol agents for plants.  
XX  
XX Disclosure; SEQ ID NO 3630; 870bp; English.  
PS  
XX The invention relates to new Proteus mirabilis polypeptides and  
CC polynucleotides. The invention also relates to antibodies against the  
CC polypeptides, methods for producing the polypeptides, a method of  
CC generating vaccines for immunising an individual against P. mirabilis, a  
CC method for evaluating a compound for the ability to bind a P. mirabilis  
CC polypeptide and a method for screening test compounds for anti-bacterial  
CC activity. The polypeptides and polynucleotides are useful as molecular  
CC targets for diagnosing, preventing and treating pathological conditions  
CC resulting from bacterial infection, as reagents for diagnosis of  
CC bacterial diseases, as components of antibacterial vaccines, as targets  
CC for antibacterial drugs or as bio-control agents for plants. This  
CC sequence represents a Proteus mirabilis polynucleotide of the invention.  
XX  
SQ Sequence 1701 BP; 560 A; 292 C; 326 G; 523 T; 0 U; 0 Other;  
  
Query Match 75.0%; Score 15; DB 10; Length 1701;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 4 TTGGTTGGCAACAC 18  
|||||  
DB 1271 TTGGTTGGCAACAC 1285  
  
RESULT 49  
ACA21214  
ID ACA21214 standard; DNA; 2154 BP.  
XX  
XX ACA21214;  
AC  
DT 19-JUN-2003 (first entry)  
XX  
XX Prokaryotic essential gene #2871.  
DE  
XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW

KW drug design; gene.  
XX  
XX Acinetobacter baumannii.  
PN  
XX WO200277183-A2.  
XX  
XX 03-OCT-2002.  
PD  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
PF  
XX  
XX 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
PA  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR P-PSDB; ABUI7344.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 9084; 1766pp; English.  
PS  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than S. aureus, S. typhimurium,  
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2154 BP; 610 A; 460 C; 488 G; 596 T; 0 U; 0 Other;  
  
Query Match 75.0%; Score 15; DB 8; Length 2154;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GCATTGGTGGCAAC 15  
|||||  
DB 860 GCATTGGTGGCAAC 874



```
RESULT 50
ADA32702
ID ADA32702 standard; DNA; 2181 BP.
XX
AC ADA32702;
XX
DT 20-NOV-2003 (first entry)
XX
DE DNA encoding Acinetobacter baumannii protein #3989.
XX
KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
KW vaccine; plant biocontrol agent.
XX
OS Acinetobacter baumannii.
XX
PN US6562958-B1.
XX
PD 13-MAY-2003.
XX
PF 04-JUN-1999; 99US-00328352.
XX
PR 09-JUN-1998; 98US-0088701P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton G, Bush D;
XX
DR WPI; 2003-576092/54.
DR P-PSDB; ADA36828.
XX
PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
PS Example; SEQ ID NO 3989; 328pp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents DNA encoding an A. baumannii
CC protein.
XX
SQ Sequence 2181 BP; 621 A; 459 C; 492 G; 609 T; 0 U; 0 Other;
Query Match 75.0%; Score 15; DB 9; Length 2181;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTTTGGTTGGCAA 15
Db 884 GCTTTGGTTGGCAA 898
Search completed: May 7, 2006, 03:33:21
Job time : 734 secs
```

THIS PAGE LEFT BLANK

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

## OM nucleic - nucleic search, using sw model

Run on: May 8, 2006, 00:50:38 ; Search time 3759 Seconds  
(without alignments)  
248.934 Million cell updates/sec

Title: US-10-671-074-172

Perfect score: 20

Sequence: 1 gcttgggtgggcaacacat 20

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size : 1

Total number of hits satisfying chosen parameters: 82156240

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gss1:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	191	8	W35284
2	20	100.0	230	1	AW496823
3	20	100.0	231	8	T85080
4	20	100.0	238	1	AU023800
5	20	100.0	240	8	H82288
6	20	100.0	255	1	BB115146
7	20	100.0	269	1	AV318369
8	20	100.0	287	1	AV338311
9	20	100.0	314	6	CF135310
10	20	100.0	316	5	BU783931
11	20	100.0	326	5	BU784200
12	20	100.0	327	1	A1216046
13	20	100.0	349	5	BX631623
14	20	100.0	353	8	W32908
15	20	100.0	355	7	CR469238
16	20	100.0	358	8	H93327
17	20	100.0	361	1	A1846721
18	20	100.0	382	1	AA035406
19	20	100.0	388	1	AW274010
20	20	100.0	393	1	AA981734
21	20	100.0	396	4	AK212379
22	20	100.0	396	6	CF3566230

23	20	100.0	397	3	BP755053
24	20	100.0	397	5	BY367902
25	20	100.0	398	4	AK211466
26	20	100.0	403	3	BI401999
27	20	100.0	404	2	BB700991
28	20	100.0	411	2	BB835762
29	20	100.0	413	6	CF135311
30	20	100.0	414	1	AA692950
31	20	100.0	414	1	AJ647314
32	20	100.0	418	1	AW825094
33	20	100.0	423	2	BG230242
34	20	100.0	425	2	BB828031
35	20	100.0	426	6	CF138380
36	20	100.0	427	1	AA893671
37	20	100.0	427	1	AA124874
38	20	100.0	433	2	BF414932
39	20	100.0	434	6	CA567706
40	20	100.0	442	1	AW912853
41	20	100.0	442	1	AA280923
42	20	100.0	448	1	AA897422
43	20	100.0	452	2	BB849701
44	20	100.0	453	7	COT15537
45	20	100.0	457	2	BI294032
46	20	100.0	461	2	BG068406
47	20	100.0	467	1	AU022008
48	20	100.0	468	1	AI373959
49	20	100.0	471	2	BB702712
50	20	100.0	472	2	BF388801
51	20	100.0	474	1	AA823039
52	20	100.0	477	1	AA547331
53	20	100.0	478	2	BE334681
54	20	100.0	486	1	AA134269
55	20	100.0	488	1	AA890151
56	20	100.0	489	1	AA922720
57	20	100.0	490	1	AI346535
58	20	100.0	490	7	CO692599
59	20	100.0	497	8	H78345
60	20	100.0	500	1	AI409296
61	20	100.0	500	2	BI181774
62	20	100.0	501	1	AW825321
63	20	100.0	501	3	BP390554
64	20	100.0	504	1	AI436585
65	20	100.0	513	7	CR773353
66	20	100.0	515	1	AI147585
67	20	100.0	515	1	AA254887
68	20	100.0	516	1	AI786162
69	20	100.0	519	1	AA054534
70	20	100.0	520	1	AA194765
71	20	100.0	531	1	AA985630
72	20	100.0	535	1	AW182749
73	20	100.0	540	2	BI295511
74	20	100.0	541	2	BG797199
75	20	100.0	551	1	AU022805
76	20	100.0	554	1	AI208073
77	20	100.0	558	1	AW985641
78	20	100.0	562	1	AU022552
79	20	100.0	569	1	AI218049
80	20	100.0	569	6	CF142596
81	20	100.0	572	3	BI555122
82	20	100.0	574	5	EU899040
83	20	100.0	579	7	CO680080
84	20	100.0	582	3	BF772955
85	20	100.0	597	2	BB633141
86	20	100.0	597	2	BE300931
87	20	100.0	601	2	BI181261
88	20	100.0	609	7	CK618770
89	20	100.0	614	3	BF768678
90	20	100.0	615	2	BI186249
91	20	100.0	638	1	AW303367
92	20	100.0	645	7	CK620335
93	20	100.0	651	7	CO602159
94	20	100.0	675	7	CO684193
95	20	100.0	677	5	BY763137

BP755053	BP755053
BY367902	BY367902
AK211466	Mus muscu
BI401999	MI-P-CP0-
BB700991	BB700991
BB835762	BB835762
CF135311	UI-HF-BR0
AA692950	vi60403.s
AJ647314	AJ647314
AW825094	u810905.y
BG230242	uv04808.x
BB828031	BB828031
CF138380	UI-HF-BN0
AA893671	EST197474
AA124874	mp73c08.f
BF414932	UI-R-BJ2
CA567706	KO417G10-
AW912853	uf45g12.y
AA280923	ea97b12.z
AA897422	al48c09.s
BB849701	uw88a11.y
COT15537	DG14-254d
BI294032	UI-R-DK0-
BG068406	H3065805-
AU022008	AU022008
AI373959	q291d11.x
BB702712	BB702712
BF388801	UI-R-BS2
AA823039	vw39h07.r
AA547331	vk71a01.s
BE334681	u882803.y
AA134269	z022g12.s
AA890151	al53f04.s
AA922720	oh91h05.s
AI346535	gp46b01.x
CO692599	DG11-50a1
H78345	yu79e12.r1
AI409296	EST237588
BI181774	UNL-P-FN-
AW825321	u813f11.y
BP390554	BP390554
AI436585	t103d08.x
CR773353	KXFZp459K
AI147585	q2b1f10.x
AA254887	mz78c05.r
AI786162	uf11f12.y
AA054534	z83d06.s
AA194765	zq06h06.s
AA985630	on44c02.s
AW182749	xp97b11.x
BI295511	UI-R-DK0-
BG797199	ic12a10.x
AU022805	AU022805
AI208073	qg46c06.x
AW985641	uf83h08.y
AU022552	AU022552
AI218049	qh29b07.x
CF142596	UI-HF-BR0
BI555122	603236356
EU899040	mai144g08.
CO680080	DG11-110e
BF772955	BB633141
BE300931	bb10d07.x
BI181261	UNL-P-FN-
CK618770	mk16c02.y
BF768678	BP768678
BI186249	UNL-P-FN-
AW303367	mv18c03.x
CK620335	x110h10.y
CO602159	DG8-235a3
CO684193	DG11-17h5
BY763137	BY763137

c 96	20	100.0	679	5	BX919154	BX919154	169	15	75.0	272	2	BB229577	BB229577
c 97	20	100.0	683	2	BE669435	7e24e11.x	c 170	15	75.0	278	1	BB044432	BB044432
c 98	20	100.0	685	2	BU632942	UI-H-DP0-	c 171	15	75.0	286	1	BQ568896	GI118a05.
c 99	20	100.0	686	5	BQ604814	MI-P-CP1-	c 172	15	75.0	289	1	AI645624	YU06h09.x
c 100	20	100.0	690	2	BE641188	BE641188	c 173	15	75.0	308	1	BB395355	BB395355
c 101	20	100.0	692	6	CA418313	UI-H-HF0-	c 174	15	75.0	316	6	CB708932	ABGNNUC:C
c 102	20	100.0	702	2	C0812298	AGENCOURT	c 175	15	75.0	360	8	CX628063	GBN007C15
c 103	20	100.0	711	2	BE850923	uw9oc06.y	c 176	15	75.0	360	8	D67487	CELK053B6F
c 104	20	100.0	740	7	CR744009	CR744009	c 177	15	75.0	369	1	AA917879	OL18f10.8
c 105	20	100.0	754	2	B1183845	UNL-P-PN-	c 178	15	75.0	375	5	BY541041	BY541041
c 106	20	100.0	763	7	CK457789	922121.MA	c 179	15	75.0	382	5	BY375407	BY375407
c 107	20	100.0	764	7	CK457893	922121.MA	c 180	15	75.0	382	5	CO993656	UMC-pd0fo
c 108	20	100.0	765	1	AW117498	xd92e10.x	c 181	15	75.0	383	5	BY481831	BY481831
c 109	20	100.0	772	1	AA746346	oa56e03.r	c 182	15	75.0	385	5	AX491612	DKF2p781G
c 110	20	100.0	787	2	B1183019	UNL-P-PN-	c 183	15	75.0	386	6	CB434508	611128.MA
c 111	20	100.0	793	1	AA727282	vu98h12.x	c 184	15	75.0	387	5	BY547976	BY547976
c 112	20	100.0	795	3	BT415030	602990433	c 185	15	75.0	390	4	AK192846	Mus.muscu
c 113	20	100.0	800	2	B1183720	UNL-P-PN-	c 186	15	75.0	393	5	EX568123	BX568123
c 114	20	100.0	801	7	C0802905	AGENCOURT	c 187	15	75.0	394	3	BP756307	BP756307
c 115	20	100.0	805	7	CJ009766	CJ009766	c 188	15	75.0	399	3	BY482043	BY482043
c 116	20	100.0	820	7	CK457570	921879.MA	c 189	15	75.0	406	1	AL713529	DKF2p686E
c 117	20	100.0	820	7	CK459307	921879.MA	c 190	15	75.0	406	1	BY541383	BY541383
c 118	20	100.0	823	2	BT233985	UNL-P-PN-	c 191	15	75.0	407	5	CF504830	USDA-PP_1
c 119	20	100.0	861	7	C0814330	AGENCOURT	c 192	15	75.0	427	1	AW054698	w89Sh03.x
c 120	20	100.0	930	2	BG110436	602278845	c 193	15	75.0	427	2	BB675201	BB675201
c 121	20	100.0	1094	4	AK039265	Mus.muscu	c 194	15	75.0	438	6	CB794341	AMGNNUC:T
c 122	20	100.0	1341	10	AY415464	AY415464.Homo.sapi	c 195	15	75.0	435	6	BY425678	BY425678
c 123	20	100.0	1341	10	AY415466	AY415466.Mus.muscu	c 196	15	75.0	437	3	BP649872	BP649872
c 124	20	100.0	2652	4	AK043583	AK043583	c 197	15	75.0	438	2	BB830772	BB830772
c 125	20	100.0	3906	4	AK053633	AK053633	c 198	15	75.0	438	10	CL328219	RFC144_24
c 126	18	90.0	250	2	BE288439	BE288439	c 199	15	75.0	440	1	AA177509	mt24f11.r
c 127	18	90.0	281	2	BE290980	BE290980	c 200	15	75.0	441	5	BY515655	BY515655
c 128	18	90.0	388	3	BM735504	MONO1.19	c 201	15	75.0	443	8	W82609	mf05h11.r1
c 129	17	85.0	193	5	BY361261	BY361261	c 202	15	75.0	449	8	CX585670	TT8000280
c 130	17	85.0	396	9	CT794625	SALK.0532	c 203	15	75.0	454	5	BY376903	BY376903
c 131	17	85.0	490	10	AL769721	Arabidops	c 204	15	75.0	454	5	BY523459	BY523459
c 132	17	85.0	613	11	FR0006731	290541.F.rubripes	c 205	15	75.0	455	5	BY486846	BY486846
c 133	17	85.0	618	11	FR0006082	289892.F.rubripes	c 206	15	75.0	458	1	AA312199	EST182886
c 134	17	85.0	619	11	FR0007302	291112.F.rubripes	c 207	15	75.0	459	5	BY571693	BY571693
c 135	17	85.0	689	10	CZ118061	CZ118061.OA.BBa000	c 208	15	75.0	467	1	AW907527	ue16h03.x
c 136	17	85.0	705	2	BG856908	BG856908.1024048G0	c 209	15	75.0	467	5	BY555789	BY555789
c 137	17	85.0	930	6	CD246789	CD246789.AGENCOURT	c 210	15	75.0	471	9	BZ259723	CH230-383
c 138	16	80.0	287	2	BB369776	BB369776	c 211	15	75.0	480	6	CF509736	USDA-PP_1
c 139	16	80.0	411	6	CF613800	CF613800.CES008101	c 212	15	75.0	486	3	BP063075	BP063075
c 140	16	80.0	443	8	CA3085	SW3ICA1093S	c 213	15	75.0	495	3	BM105218	508622.MA
c 141	16	80.0	466	6	CF613858	CF613858.CES008176	c 214	15	75.0	498	3	BP002665	BP002665
c 142	16	80.0	475	1	AW521319	UI-R-B00-	c 215	15	75.0	510	5	BU457032	603772243
c 143	16	80.0	497	6	CA670500	w18u1.pk0	c 216	15	75.0	512	1	AA601404	no16f10.8
c 144	16	80.0	516	5	BQ590451	BQ590451.E012840-0	c 217	15	75.0	519	9	AQ727430	HS.5450.A
c 145	16	80.0	577	5	BQ510918	BQ510918.ESF618333	c 218	15	75.0	533	6	CD621513	56005596H
c 146	16	80.0	610	5	BQ448043	BQ448043.UI-H-EU1-	c 219	15	75.0	538	2	BG092431	mac11a02.
c 147	16	80.0	616	1	AA154353	AA154353.m814h07.r	c 220	15	75.0	541	3	BP103230	BP103230
c 148	16	80.0	620	1	AI172179	EST218174	c 221	15	75.0	546	5	EX514163	EX514163
c 149	16	80.0	620	2	BF504327	BF504327.AT05668.5	c 222	15	75.0	551	9	AQ541328	RPCI-11-3
c 150	16	80.0	622	9	AQ731279	AQ731279.HS.5517.B	c 223	15	75.0	553	3	BP956092	BP956092
c 151	16	80.0	622	10	CL749110	CL749110.OR.BBa011	c 224	15	75.0	554	6	CD621511	56005588H
c 152	16	80.0	631	6	CF615152	CF615152.CES008984	c 225	15	75.0	555	2	BE873236	601451771
c 153	16	80.0	633	9	AQ090652	AQ090652.RFC01457	c 226	15	75.0	561	1	AV601167	AV601167
c 154	16	80.0	645	9	AQ327286	AQ327286.nbx0040P	c 227	15	75.0	562	6	CA950956	ir32g08.y
c 155	16	80.0	670	5	BY171757	BY171757	c 228	15	75.0	566	3	BQ142569	Cont1856
c 156	16	80.0	688	5	BW283929	BW283929	c 229	15	75.0	566	6	CB427376	CB427376
c 157	16	80.0	849	10	CG089045	CG089045.PUJFN22TD	c 230	15	75.0	569	10	CZ868210	OC_Ba022
c 158	16	80.0	895	10	AG043961	AG043961.Pan.trogl	c 231	15	75.0	575	1	CZ839323	OC_Ba026
c 159	16	80.0	975	11	CNS041A7	AL269944.Tetraodon	c 232	15	75.0	577	9	BZ346712	hm33d06.b
c 160	15	75.0	187	7	C0312778	C0312778	c 233	15	75.0	579	7	CK902479	CK902479
c 161	15	75.0	212	1	AA184751	AA184751.mt94e12.r	c 234	15	75.0	584	6	CD621509	56005580H
c 162	15	75.0	224	2	BE287850	BE287850	c 235	15	75.0	593	6	CD621507	56005572H
c 163	15	75.0	232	9	AQ036093	AQ036093.CIT-HGP-2	c 236	15	75.0	595	7	CK376612	1ah87h05.
c 164	15	75.0	235	1	BB326112	BB326112	c 237	15	75.0	600	10	CL322869	RPCI44.43
c 165	15	75.0	251	1	AV236517	AV236517	c 238	15	75.0	607	7	CV129240	MdFrt3091
c 166	15	75.0	258	1	BB043559	BB043559	c 239	15	75.0	607	7	CV573233	od33h02.y
c 167	15	75.0	266	9	AO677203	HS_5514.A	c 240	15	75.0	608	7	CV130092	MdFrt3087
c 168	15	75.0	269	1	AV325659	AV325659	c 241	15	75.0	610	3	BP758612	BP758612

C 242	15	75.0	610	7	CV572195	CV572195 od2a11.Y	315	15	75.0	957	11	DQ040538
C 243	15	75.0	612	3	BP002831	BP002831 BP002831	C 316	15	75.0	964	3	BQ229478
C 244	15	75.0	613	2	BE622702	BE622702 BE622702	C 317	15	75.0	968	5	BUS27695
C 245	15	75.0	614	7	CR283230	CR283230 CR283230	318	15	75.0	971	5	BQ682739
C 246	15	75.0	624	6	CB422065	CB422065 595134 MA	319	15	75.0	976	10	CL127403
C 247	15	75.0	639	9	BH834387	BH834387 BACP6-L0	320	15	75.0	981	10	CL127470
C 248	15	75.0	643	5	BW270890	BW270890 BW270890	321	15	75.0	997	1	AL569480
C 249	15	75.0	647	9	BH835070	BH835070 BACP7-B0	C 322	15	75.0	1002	10	CZ951230
C 250	15	75.0	648	10	AG966736	AG966736 Droepophil	C 323	15	75.0	1023	3	BM460435
C 251	15	75.0	649	10	C2637518	C2637518 OM_Ba018	C 324	15	75.0	1027	8	DN575431
C 252	15	75.0	650	6	CB445030	CB445030 696270 MA	325	15	75.0	1029	6	CA478752
C 253	15	75.0	655	2	BF168549	BF168549 601775451	C 326	15	75.0	1069	2	BE963867
C 254	15	75.0	660	9	BZ750227	BZ750227 PUDBU3RTB	C 327	15	75.0	1095	10	AV404703
C 255	15	75.0	660	10	CW366570	CW366570 fdbb001f0	C 328	15	75.0	1111	9	CC187166
C 256	15	75.0	664	3	BW179336	BW179336 UI-B-E00-	C 329	15	75.0	1114	8	DR144960
C 257	15	75.0	664	10	CW205958	CW205958 104_634_1	C 330	15	75.0	1171	2	BE420996
C 258	15	75.0	670	6	CB427738	CB427738 603334 MA	C 331	15	75.0	1217	8	DR144424
C 259	15	75.0	678	10	CZ771597	CZ771597 OC_Ba012	332	15	75.0	1218	8	DN804982
C 260	15	75.0	683	7	CN407574	CN407574 170006001	333	15	75.0	1238	6	CD495465
C 261	15	75.0	690	5	BU072133	BU072133 im54f08.Y	334	15	75.0	1313	3	BI562427
C 262	15	75.0	707	9	CC174377	CC174377 ZMBBEC029	C 335	15	75.0	1333	8	DN694526
C 263	15	75.0	709	6	CB444693	CB444693 695986 MA	C 336	15	75.0	1508	4	CR621279
C 264	15	75.0	712	6	CB422415	CB422415 595518 MA	C 337	15	75.0	1616	4	AK017910
C 265	15	75.0	720	10	CW622471	CW622471 OP_Ba002	C 338	15	75.0	2987	4	AK037533
C 266	15	75.0	721	7	CK223023	CK223023 704398777	339	15	75.0	4386	10	AY405198
C 267	15	75.0	726	5	BW433491	BW433491 BW433491	340	15	75.0	4386	10	AY405198
C 268	15	75.0	727	10	CW486655	CW486655 fdbb001f2	C 341	14	70.0	109	10	AX891486
C 269	15	75.0	732	1	AJ819991	AJ819991 AJ819991	342	14	70.0	126	7	CK549660
C 270	15	75.0	733	1	AJ819990	AJ819990 AJ819990	343	14	70.0	148	2	BI278318
C 271	15	75.0	734	5	BU104194	BU104194 60306557	344	14	70.0	180	1	AV423825
C 272	15	75.0	737	8	CX6411381	CX6411381 UCRPT02_5	C 345	14	70.0	188	9	AZ845817
C 273	15	75.0	742	10	CZ846572	CZ846572 OC_Ba023	C 346	14	70.0	196	1	AW401624
C 274	15	75.0	755	2	BG351150	BG351150 106E04 MA	C 347	14	70.0	196	3	BNM001994
C 275	15	75.0	761	7	CK364433	CK364433 AGENCOURT	C 348	14	70.0	208	9	CE214869
C 276	15	75.0	764	7	CV477581	CV477581 57485.1 D	349	14	70.0	209	9	AZ785736
C 277	15	75.0	766	10	CQ350731	CQ350731 OG1EF45TH	350	14	70.0	215	9	BH778624
C 278	15	75.0	776	1	BB023057	BB023057 BB023057	C 351	14	70.0	220	7	CF740719
C 279	15	75.0	777	10	CZ027462	CZ027462 OM_Ba000	C 352	14	70.0	242	1	AJ601563
C 280	15	75.0	778	8	DN514011	DN514011 1251033 M	353	14	70.0	242	7	CO873533
C 281	15	75.0	780	10	CW389454	CW389454 fdbb001f0	354	14	70.0	267	2	BI276869
C 282	15	75.0	784	2	BG212399	BG212399 RST31986	C 355	14	70.0	277	1	AA833295
C 283	15	75.0	786	10	CZ763668	CZ763668 OC_Ba011	356	14	70.0	291	1	AV050281
C 284	15	75.0	790	9	BZ771311	BZ771311 mcb75a02.	C 357	14	70.0	297	2	BE500706
C 285	15	75.0	795	2	B1115894	B1115894 602866664	358	14	70.0	299	1	AI264523
C 286	15	75.0	795	2	BE3866290	BE3866290 601274543	359	14	70.0	307	1	AW183818
C 287	15	75.0	795	9	BH269934	BH269934 CH230-6M1	360	14	70.0	312	2	BE064238
C 288	15	75.0	799	7	CO483931	CO483931 GOQ205_B3	C 361	14	70.0	317	2	BE210876
C 289	15	75.0	800	7	CO888130	CO888130 BvGen.16	C 362	14	70.0	319	6	CF034569
C 290	15	75.0	804	8	CX636290	CX636290 UCRPT02_1	C 363	14	70.0	319	10	CW736579
C 291	15	75.0	829	9	CC072519	CC072519 CSU-K33r.	C 364	14	70.0	322	8	CX110573
C 292	15	75.0	835	6	CD247662	CD247662 AGENCOURT	365	14	70.0	330	1	AV815303
C 293	15	75.0	849	6	CF690239	CF690239 CCACZ64TO	C 366	14	70.0	334	3	BQ242966
C 294	15	75.0	851	10	CG821415	CG821415 SOYEU94TH	367	14	70.0	341	2	BF287647
C 295	15	75.0	853	9	CZ727604	CZ727604 OC_Ba005	C 368	14	70.0	355	5	BY000940
C 296	15	75.0	854	10	CC5223854	CC5223854 CH240_372	C 369	14	70.0	359	8	Z36258
C 297	15	75.0	859	6	CF716836	CF716836 CCAB844TF	C 370	14	70.0	360	5	BU571700
C 298	15	75.0	865	6	CB209520	CB209520 AGENCOURT	C 371	14	70.0	366	6	CA713251
C 299	15	75.0	871	10	CW781203	CW781203 OP_Ba008	372	14	70.0	368	3	BI429845
C 300	15	75.0	882	6	CF066469	CF066469 OC_Ba028	C 373	14	70.0	368	10	CE527995
C 301	15	75.0	883	6	CF066458	CF066458 Ac424_Amp	374	14	70.0	373	1	AW697493
C 302	15	75.0	884	8	DN583625	DN583625 90893580	C 375	14	70.0	373	7	CN210184
C 303	15	75.0	887	9	BZ240426	BZ240426 CH230-507	C 376	14	70.0	376	5	BY362535
C 304	15	75.0	891	2	BE974610	BE974610 601680615	C 377	14	70.0	376	5	CO5745
C 305	15	75.0	897	10	DU056016	DU056016 94353 Tom	C 378	14	70.0	377	1	AW415843
C 306	15	75.0	907	2	BF673821	BF673821 602135971	379	14	70.0	377	10	CZ2113192
C 307	15	75.0	907	10	CG771672	CG771672 TCB48_4_C	C 380	14	70.0	378	8	DN186872
C 308	15	75.0	913	5	BQ892605	BQ892605 AGENCOURT	C 381	14	70.0	382	1	AA835399
C 309	15	75.0	922	2	BE699925	BE699925 602127380	C 382	14	70.0	383	2	BE061335
C 310	15	75.0	931	9	BZ153865	BZ153865 CH230-300	C 383	14	70.0	389	2	BE488034
C 311	15	75.0	931	9	BZ771343	BZ771343 mcb76a02.	C 384	14	70.0	397	8	D24184
C 312	15	75.0	932	5	BU534072	BU534072 AGENCOURT	C 385	14	70.0	400	2	BB480612
C 313	15	75.0	941	2	BF784012	BF784012 602107856	C 386	14	70.0	402	2	BF899104
C 314	15	75.0	945	10	CL127356	CL127356 ISB1-92F2	C 387	14	70.0	402	3	BU243532

DQ040538	Pan trogl
BQ229478	AGENCOURT
BUS27695	AGENCOURT
BQ682739	AGENCOURT
CL127403	ISB1-92H1
CL127470	ISB1-9211
AL569480	AL569480
CZ951230	262076 To
BM460435	AGENCOURT
DN575431	91953185
CA478752	AGENCOURT
BE963867	601657635
AV404703	Homo sapi
CC187166	CH261-91M
DR144960	49075243
BE420996	HMM004.G0
DR144424	49135120
DN804982	76947221
CD495465	CD416-G01
BI562427	603255590
DN694526	CGX94-G01
CR621279	full-length
AK017910	Mus muscu
AK037533	Mus muscu
AY405198	Homo sapi
AY405199	Pan trogl
AX891486	Arabidops
CK549660	swkx0_006
BI278318	UI-R-CWO-
AV423825	AV423825
AZ845817	2M0145113
AW401624	UI-HF-BKO
BNM001994	103109991
CE214869	tigr-ges-
AZ785736	2M0029J20
BH778624	fzmb0013f0
CF740719	CR740719
AJ601563	AJ601563
BOVgen_01	BOVgen_01
BI276869	UI-R-CXO-
AA833295	u05d07_r
AV050281	AV050281
BE500706	WHE0991-0
AI264523	q139n08.x
AW183818	xJ88c06.x
BE064238	CM4-BT030
BE210876	so54f12.y
CF034569	OCF6h05.Y
CW736579	MARC_8256
CX110573	EIO50F09
AV815303	AV815303
BQ242966	TaeI5020H
BF287647	EST452238
CA713251	wdk3c.pk0
BI429845	fq75g07.x
CE527995	tigr-ges-
AW697493	ST61B12.P
CN210184	4116187.B
BY362535	BY362535
CO5745	CO5745 Huma
AW415843	50391.MAR
CZ2113192	AI1AA-aaf4
DN186872	HO28B04w
AA835399	ak71911.B
BE061335	IL0-BT016
BE488034	177262.BA
D24184	RICR1494.R
BB480612	165748.BA
BF899104	CM2-MT018
BU243532	BU243532

388	14	70.0	402	9	AQ045047	RPC111-34	C 461	14	70.0	533	10	CW333392	CW333392 104 831_1
389	14	70.0	403	2	A2439012	1M0229J21	462	14	70.0	534	7	CO070776	CO070776 GR_Ea27P
390	14	70.0	405	2	BF613246	de31f10.Y	463	14	70.0	535	7	CO077973	CO077973 GR_Ea40A
391	14	70.0	405	9	AQ210223	HS_3229.A	C 464	14	70.0	536	1	AU248157	AU248157 AU248157
392	14	70.0	409	7	CO184873	EC28716.5	C 465	14	70.0	536	1	AZ882244	AZ882244 RPCI-23-1
393	14	70.0	409	9	B2874641	CH240_240	466	14	70.0	538	5	BW334329	BW334329 BW334329
394	14	70.0	410	6	CF046694	QKC28h12.	467	14	70.0	540	7	CN409294	CN409294 170006001
395	14	70.0	412	9	AQ952216	Sheared.D	C 468	14	70.0	545	1	AL809237	AL809237 AL809237
396	14	70.0	413	1	AV768715	AV768715	C 469	14	70.0	545	2	BE517452	BE517452 WHR0626.B
397	14	70.0	417	9	AQ728310	HS_5446.B	C 470	14	70.0	545	9	AQ637851	AQ637851 927P1-18H
398	14	70.0	420	5	EU400403	EU400403	C 471	14	70.0	547	2	BG545941	BG545941 602573236
399	14	70.0	421	9	AQ462784	HS_5203.A	472	14	70.0	547	7	CO998058	CO998058 pam01-17m
400	14	70.0	424	1	A1116062	A1116062 uc14h08.r	473	14	70.0	548	2	BG040887	BG040887 NXSI_116
401	14	70.0	426	1	AV415326	AV415326	474	14	70.0	548	7	CN409295	CN409295 170005321
402	14	70.0	427	1	AT665154	6V5007804	475	14	70.0	550	2	BE341206	BE341206 EST0345307
403	14	70.0	427	1	AV802341	AV802341	C 476	14	70.0	550	6	CD866685	CD866685 AZ02.104C
404	14	70.0	430	1	AA789928	vx70d09.r	C 477	14	70.0	550	6	CF037882	CF037882 QCH13905.
405	14	70.0	433	6	CD393170	Gm_CK1264	C 478	14	70.0	550	6	DT048239	DT048239 COT_CY_B0
406	14	70.0	434	9	BH785706	fzmb013f0	479	14	70.0	550	8	DT052233	DT052233 COT_EV_E0
407	14	70.0	437	7	CN246034	EST011919	480	14	70.0	551	3	BP241503	BP241503 BP241503
408	14	70.0	437	7	CN446278	CN446278 CM390r.La	C 481	14	70.0	551	5	BU439513	BU439513 604143961
409	14	70.0	440	1	AV766417	AV766417	C 482	14	70.0	553	8	DR708435	DR708435 Asm_09302
410	14	70.0	442	3	EM895373	EM895373	C 483	14	70.0	554	6	CD887570	CD887570 G118.105K
411	14	70.0	443	8	R83983	R83983 15942 Lambd	484	14	70.0	556	3	BM092380	BM092380 sah12e05.
412	14	70.0	451	9	AQ667768	HS_2137.B	485	14	70.0	556	8	DT107557	DT107557 JGI_ANN8
413	14	70.0	453	7	CO100316	CO100316 GR_Ea25K	486	14	70.0	556	9	AZ829391	AZ829391 2M0107101
414	14	70.0	455	8	T89594	vd398a10.e1	C 487	14	70.0	557	9	AQ623755	AQ623755 HS_5319.A
415	14	70.0	456	6	CD723501	CD723501 4036623.1	C 488	14	70.0	562	1	AW131487	AW131487 xf30d01.x
416	14	70.0	461	7	CV324723	CM4_F7010	489	14	70.0	562	2	BG723175	BG723175 602690766
417	14	70.0	466	3	EM447165	DSAA06G06	490	14	70.0	563	5	BM165766	BM165766 BW165766
418	14	70.0	466	7	CF848723	CF848723 p8MA008xA	491	14	70.0	565	3	BM895247	BM895247 952070H08
419	14	70.0	467	9	AQ211502	HS_3243.A	492	14	70.0	566	7	CO073255	CO073255 GR_Ea32M
420	14	70.0	469	3	B0117231	EST02807	493	14	70.0	567	6	CD396029	CD396029 Gm_Ck1623
421	14	70.0	469	5	BU807219	haa13f04.	494	14	70.0	568	1	AV615349	AV615349 AV615349
422	14	70.0	469	10	CL897843	ab946e06.	495	14	70.0	572	1	AW973043	AW973043 EST385140
423	14	70.0	474	3	BM967955	LM24HW008	C 496	14	70.0	573	8	W15870	W15870 mb55a12.r1
424	14	70.0	474	6	CA704214	wk1c.pk0	C 497	14	70.0	574	1	AJ635934	AJ635934 AJ635934
425	14	70.0	474	7	CV610752	zcm28b04.	C 498	14	70.0	574	6	CA626630	CA626630 w1ln.pk01
426	14	70.0	478	3	BJ222411	BJ222411	C 499	14	70.0	575	3	BP006388	BP006388 BP006388
427	14	70.0	480	1	AJ469257	AJ469257	C 500	14	70.0	576	1	AA521736	AA521736 vi15f12.r
428	14	70.0	486	5	BU088613	BU088613 Na_L3_46D	C 501	14	70.0	581	5	BQ462228	BQ462228 HD01P167
429	14	70.0	488	7	CO086103	GR_Ea03N	502	14	70.0	582	3	BP282426	BP282426 BP282426
430	14	70.0	490	3	BM345235	rf53b10.Y	503	14	70.0	582	3	BP282458	BP282458 BP282458
431	14	70.0	493	1	AU124061	AU124061	504	14	70.0	582	3	BP283061	BP283061 BP283061
432	14	70.0	494	5	BX672761	BX672761	505	14	70.0	582	3	BP283149	BP283149 BP283149
433	14	70.0	495	3	BM324195	PIC1_25.E	506	14	70.0	582	3	BP283678	BP283678 BP283678
434	14	70.0	499	10	CW097535	104_463_1	507	14	70.0	582	3	BP363148	BP363148 BP363148
435	14	70.0	500	1	AW074134	xb08c05.x	C 508	14	70.0	582	6	CD912426	CD912426 G550.114G
436	14	70.0	500	3	BP394524	BP394524	509	14	70.0	583	3	BP351459	BP351459 BP351459
437	14	70.0	501	2	AW500065	UI-HF-BN0	510	14	70.0	585	2	BI285563	BI285563 UI-R-CW08
438	14	70.0	502	2	BG790687	aae75a05.	C 511	14	70.0	585	5	C23111	C23111 C23111 Japa
439	14	70.0	502	5	BY380060	BY380060	C 512	14	70.0	586	8	DN887408	DN887408 nag04907.
440	14	70.0	502	9	AQ519958	HS_5185.B	C 513	14	70.0	586	11	CR140186	CR140186 Reverse.s
441	14	70.0	504	10	CZ568763	C023ALL-F	514	14	70.0	587	3	BP364999	BP364999 BP364999
442	14	70.0	505	6	CB829921	ri07a02.Y	515	14	70.0	588	3	BP350472	BP350472 BP350472
443	14	70.0	507	11	DE070674	Ocyziaa1	C 516	14	70.0	588	6	CB636357	CB636357 OSJNEA01M
444	14	70.0	509	6	CA777302	ip05h07.Y	517	14	70.0	588	7	CK010173	CK010173 29365x8iC
445	14	70.0	510	2	BF624269	BF624269 HVSM8A001	518	14	70.0	589	9	AZ840622	AZ840622 2M0138E14
446	14	70.0	511	1	AA017621	AA017621 ze37G04.r	519	14	70.0	589	2	BG357233	BG357233 OV2_11.D1
447	14	70.0	512	3	BI751078	BI751078 Tau1_05e0	C 520	14	70.0	593	10	CW129985	CW129985 104_512_1
448	14	70.0	512	4	AK196189	AK196189 Mus_muscu	521	14	70.0	593	3	BM062755	BM062755 KS01046H0
449	14	70.0	513	7	CN547433	CN547433 EST_15422	C 522	14	70.0	594	6	CB348301	CB348301 CAB2SG000
450	14	70.0	517	6	CB188733	CB188733 k620c03.Y	C 523	14	70.0	595	3	BJ612792	BJ612792 BJ612792
451	14	70.0	519	5	BX951008	BX951008 DKFZp781P	524	14	70.0	595	9	AZ859980	AZ859980 2M0165114
452	14	70.0	519	7	CN203124	Tor3361.G	C 525	14	70.0	595	10	CE342219	CE342219 t1gr-g88s-
453	14	70.0	520	1	AV771243	AV771243	C 526	14	70.0	596	3	BM100498	BM100498 EBna01.SS
454	14	70.0	520	2	BE422591	WHR0055.B	C 527	14	70.0	596	5	CA014595	CA014595 HT11M15r
455	14	70.0	522	9	AQ539181	RPCI-11-3	C 528	14	70.0	597	5	BW331066	BW331066 BW331066
456	14	70.0	524	3	BP692640	BP692640	C 529	14	70.0	597	10	CW736759	CW736759 MARC_8259
457	14	70.0	526	1	AW942989	AW942989	C 530	14	70.0	598	11	PT018E03R	PT018E03R Parametciu
458	14	70.0	526	9	AQ784262	HS_3087.B	C 531	14	70.0	600	2	BG487684	BG487684 FMI_73_A0
459	14	70.0	527	6	CA642568	wre1n.pk0	532	14	70.0	600	2	BE365842	BE365842 F11_2.F11
460	14	70.0	532	1	AW720424	LjNESt22g	C 533	14	70.0	600	5	BW202619	BW202619 BW202619

534	14	70.0	600	7	CO070557	CO070557 GR_Ba27K	607	14	70.0	656	9	BH119617	BH119617 RPCI-24-2
535	14	70.0	600	10	AG924751	AG924751 Drosophila	608	14	70.0	656	9	CC936071	CC936071 ZMMB5019
536	14	70.0	601	8	DN146324	DN146324 4844_A05	609	14	70.0	656	10	CW333393	CW333393 104_831_1
537	14	70.0	602	6	CA643625	CA643625 wreln.pk0	610	14	70.0	657	3	CD414634	CD414634 Gm_CK4683
538	14	70.0	605	6	CA067996	CA067996 SCQAD105	c 611	14	70.0	659	3	BJ269922	BJ269922 BJ269922
539	14	70.0	605	7	CV472874	CV472874 21027_1.D	c 612	14	70.0	660	6	CB517916	CB517916 chslcgb54
540	14	70.0	605	9	AQ390942	AQ390942 CITBI-E1	c 613	14	70.0	660	9	BZ120814	BZ120814 ch230-423
541	14	70.0	605	11	FR0031892	AL028261 Fugu rubr	614	14	70.0	660	10	CB553594	CB553594 tigr-g88-
542	14	70.0	606	6	CA067993	CA067993 SCQAD105	615	14	70.0	661	5	BW325447	BW325447 BW325447
543	14	70.0	607	3	BU555526	BU555526 BU555526	616	14	70.0	661	6	CB851363	CB851363 UI-CF-DU1
544	14	70.0	607	5	BU339151	BU339151 603515490	c 617	14	70.0	665	5	CD621505	CD621505 55049374H
545	14	70.0	608	6	CB348376	CB348376 CAB28G000	c 618	14	70.0	666	5	BQ806619	BQ806619 WHB3581.B
546	14	70.0	608	10	CZ561754	CZ561754 1_p.bw014	619	14	70.0	667	3	BI751087	BI751087 Ta01_05Fo
547	14	70.0	609	8	CV880356	CV880356 Mdel6018k	c 620	14	70.0	670	6	CA088550	CA088550 SCRFAM212
548	14	70.0	610	1	AW447901	AW447901 BRY_1017	c 621	14	70.0	670	10	CZ034832	CZ034832 OM_Ba001
549	14	70.0	610	3	BM895264	BM895264 952073A11	622	14	70.0	671	8	DN991406	DN991406 TC118874
550	14	70.0	610	5	BQ605548	BQ605548 BRY_1017	623	14	70.0	674	7	CO105998	CO105998 GR_B0003
551	14	70.0	612	6	CA227125	CA227125 SCVPLP304	624	14	70.0	674	8	DT0337845	DT0337845 VLI132D10
552	14	70.0	616	2	BG717631	BG717631 602698252	c 625	14	70.0	674	9	BH925458	BH925458 odi33a08
553	14	70.0	617	5	BU984301	BU984301 HF03121r	626	14	70.0	674	9	BZ193790	BZ193790 CH230-331
554	14	70.0	617	3	BQ245698	BQ245698 TaE5021C	c 627	14	70.0	676	8	CX513722	CX513722 JGI_XZG58
555	14	70.0	618	6	CB843312	CB843312 M158-3802	628	14	70.0	677	3	BI688660	BI688660 603311282
556	14	70.0	619	11	CR212154	CR212154 Reverse s	c 629	14	70.0	677	9	BZ537732	BZ537732 OCAJH77TF
557	14	70.0	620	11	CR344515	CR344515 mte1-78K1	c 630	14	70.0	677	10	CW137227	CW137227 104_523_1
558	14	70.0	621	10	CB722521	CB722521 tigr-g88-	c 631	14	70.0	678	7	CV055826	CV055826 BNEL12313
559	14	70.0	622	5	BU977459	BU977459 HAL1H12r	c 632	14	70.0	679	10	AG092171	AG092171 Pan trogl
560	14	70.0	622	5	BW324356	BW324356 BW324356	c 633	14	70.0	679	11	DE065979	DE065979 Oryzias l
561	14	70.0	624	9	CS641915	CS641915 OGUJY11TH	c 634	14	70.0	681	1	AV967402	AV967402 AV967402
562	14	70.0	624	5	BY721465	BY721465 BY721465	c 635	14	70.0	681	5	BQ766985	BQ766985 EBR008_SQ
563	14	70.0	624	6	CA165048	CA165048 SCBR2312	c 636	14	70.0	681	6	CA839554	CA839554 MCT028G11
564	14	70.0	626	5	BU401442	BU401442 603485006	c 637	14	70.0	681	6	CF587815	CF587815 USDA-PP_1
565	14	70.0	626	10	CL938514	CL938514 OA_ABA005	c 638	14	70.0	682	6	CA289466	CA289466 SCAGFL800
566	14	70.0	628	6	CB861000	CB861000 RH03111w	c 639	14	70.0	682	7	CH062373	CH062373 59970r81c
567	14	70.0	628	10	BX145114	BX145114 Danio rer	c 640	14	70.0	685	9	BH972740	BH972740 odj_46B01
568	14	70.0	629	2	BG721559	BG721559 602695167	c 641	14	70.0	686	5	BW372086	BW372086 BW372086
569	14	70.0	629	7	CV057382	CV057382 BNEL27b8	c 642	14	70.0	686	7	CV055855	CV055855 BNEL124c1
570	14	70.0	631	1	AA980353	AA980353 ua52b03_r	c 643	14	70.0	686	8	CX166023	CX166023 HESC2_35
571	14	70.0	632	1	AV405220	AV405220 AV405220	c 644	14	70.0	689	10	CG811493	CG811493 FSAAT54TR
572	14	70.0	632	3	BQ245602	BQ245602 TaE15023G	c 645	14	70.0	689	10	CZ064449	CZ064449 OM_Ba006
573	14	70.0	635	5	BW321657	BW321657 BW321657	c 646	14	70.0	690	2	BI210511	BI210511 EST528551
574	14	70.0	635	5	BW038196	BW038196 BW038196	c 647	14	70.0	690	9	CC579876	CC579876 CH240_374
575	14	70.0	636	6	CD403362	CD403362 Gm_CK2609	c 648	14	70.0	691	3	BQ247112	BQ247112 TaE15001C
576	14	70.0	636	10	CZ123223	CZ123223 OA_BBA001	c 649	14	70.0	691	3	CF201623	CF201623 RR890915N
577	14	70.0	637	5	BW199716	CE380988 tigr-g88-	c 650	14	70.0	691	8	CX899976	CX899976 JGI_CAAW8
578	14	70.0	637	5	CA087970	BW199716 BW199716	c 651	14	70.0	691	8	DN227997	DN227997 MEST1213
579	14	70.0	637	7	CA087970	CA087970 SCQAM210	c 652	14	70.0	691	10	AG296053	AG296053 Mus muscu
580	14	70.0	637	7	CK547630	CK547630 swk20_001	c 653	14	70.0	693	2	BE601684	BE601684 HVSMBH009
581	14	70.0	639	1	AJ432029	AJ432029 AJ432029	c 654	14	70.0	693	7	CN832627	CN832627 AGENCOURT
582	14	70.0	640	1	AL857957	AL857957 AL857957	c 655	14	70.0	694	9	BH924837	BH924837 odi56d12
583	14	70.0	640	7	CV057374	CV057374 BNEL27b10	c 656	14	70.0	695	5	BX484530	BX484530 DKFP2686G
584	14	70.0	641	5	CA009163	CA009163 HUI3F18r	c 657	14	70.0	695	6	CA115556	CA115556 SCBLB103
585	14	70.0	641	6	CB888641	CB888641 G118.108L	c 658	14	70.0	696	6	CA198237	CA198237 SCRFAM110
586	14	70.0	642	6	CF005959	CF005959 QB116f09	c 659	14	70.0	696	10	CG165902	CG165902 PUF2X07TB
587	14	70.0	643	3	BI602822	BI602822 603247182	c 660	14	70.0	696	10	CL180264	CL180264 104_390_1
588	14	70.0	643	6	CA215051	CA215051 SCRAD112	c 661	14	70.0	697	7	CN287516	CN287516 170005326
589	14	70.0	643	8	DR087825	DR087825 CAST100F1	c 662	14	70.0	698	5	BW294504	BW294504 BW294504
590	14	70.0	644	1	AV949479	AV949479 AV949479	c 663	14	70.0	699	1	AX895105	AX895105 AV895105
591	14	70.0	645	5	BX878778	BX878778 BX878778	c 664	14	70.0	699	7	CK066977	CK066977 72904r81c
592	14	70.0	645	10	CZ603783	CZ603783 OM_Ba012	c 665	14	70.0	700	7	CV054070	CV054070 BNEL106d4
593	14	70.0	646	5	BW325542	BW325542 BW325542	c 666	14	70.0	701	6	CF201680	CF201680 RR890915N
594	14	70.0	646	6	CD403101	CD403101 Gm_CK2575	c 667	14	70.0	706	2	BE601912	BE601912 HVSMBH010
595	14	70.0	647	5	BW251038	BW251038 BW251038	c 668	14	70.0	706	7	CO985202	CO985202 GM89023B1
596	14	70.0	648	5	BW330520	BW330520 BW330520	c 669	14	70.0	707	6	CD866766	CD866766 AZO2_104F
597	14	70.0	650	5	BQ471619	BQ471619 HV02P12r	c 670	14	70.0	707	6	CF481796	CF481796 POLI_74_E
598	14	70.0	650	6	CA210410	CA210410 SCEPSB112	c 671	14	70.0	707	8	DT085321	DT085321 JGI_ANNK3
599	14	70.0	650	6	CA732290	CA732290 wlpic.pk0	c 672	14	70.0	709	6	CA763250	CA763250 BR06008A
600	14	70.0	650	10	CZ231128	CZ231128 A1AA-baf9	c 673	14	70.0	709	7	CV053850	CV053850 BNEL103H1
601	14	70.0	651	6	CA388198	CA388198 670351_NC	c 674	14	70.0	709	9	CF060466	CF060466 POLI_65_F
602	14	70.0	652	7	CV060466	CV060466 BNEL58d7	c 675	14	70.0	711	6	CF480441	CF480441 WHB3582_A
603	14	70.0	653	10	AG136288	AG136288 Pan trogl	c 676	14	70.0	716	9	AZ908418	AZ908418 RPCI-24-2
604	14	70.0	654	5	BU991957	BU991957 DR08G17r	c 677	14	70.0	716	10	AG393890	AG393890 Mus muscu
605	14	70.0	655	5	BW234188	BW234188 BW234188	c 678	14	70.0	717	8	CX997039	CX997039 JGI_CAAQ6
606	14	70.0	655	5	BW234188	BW234188 BW234188	679	14	70.0				

680	14	70.0	717	9	CE038827	753	14	70.0	789	6	CB180222	CB180222	AGENCOURT
c 681	14	70.0	718	8	CX910314	c 754	14	70.0	790	8	CX899562	CX899562	JGI_CAAAN2
c 682	14	70.0	719	11	CR053056	c 755	14	70.0	791	1	AJ612715	AJ612715	AJ612715
683	14	70.0	719	5	BW038201	c 756	14	70.0	793	10	BX158897	BX158897	Danio rer
684	14	70.0	720	7	CV705629	c 757	14	70.0	794	6	CD048804	CD048804	AGENCOURT
c 685	14	70.0	720	7	CV705629	c 757	14	70.0	794	6	CD048804	CD048804	AGENCOURT
c 686	14	70.0	722	8	CX997040	c 758	14	70.0	795	7	CK317001	CK317001	SB02020A2
c 687	14	70.0	722	8	DT085786	c 759	14	70.0	796	9	DR857919	JGI_CARG3	PJF0A61TD
c 688	14	70.0	725	8	CX484348	c 760	14	70.0	796	8	DR857919	JGI_CARG3	PJF0A61TD
c 689	14	70.0	726	10	AG215665	c 761	14	70.0	796	9	DR857919	JGI_CARG3	PJF0A61TD
c 690	14	70.0	728	1	AA762751	c 762	14	70.0	797	2	BG344447	HVSNEG000	
c 691	14	70.0	728	5	BX914879	c 763	14	70.0	797	2	BG344447	HVSNEG000	
692	14	70.0	728	5	BY710913	c 764	14	70.0	797	2	BG344447	HVSNEG000	
693	14	70.0	728	8	DN507301	c 765	14	70.0	798	10	CW721892	BI207408	EST525448
694	14	70.0	729	3	BI560837	c 766	14	70.0	798	10	CW721892	BI207408	EST525448
c 695	14	70.0	729	10	CW462017	c 767	14	70.0	800	6	CD779190	ESY650551	
c 696	14	70.0	731	9	BZ687843	c 768	14	70.0	800	6	CD779190	ESY650551	
c 697	14	70.0	731	10	CW517842	c 769	14	70.0	800	6	CD779190	ESY650551	
c 698	14	70.0	732	5	BW331182	c 770	14	70.0	801	10	BX135746	BX135746	Danio rer
699	14	70.0	732	5	BW331182	c 771	14	70.0	801	10	BX135746	BX135746	Danio rer
c 700	14	70.0	733	2	BE277305	c 772	14	70.0	801	6	CB897606	tricol11xn	
c 701	14	70.0	733	5	BW246179	c 773	14	70.0	801	6	CB897606	tricol11xn	
c 702	14	70.0	735	2	BF674397	c 774	14	70.0	802	7	CN759569	ID0AAA25D	
c 703	14	70.0	735	6	CA224869	c 775	14	70.0	802	7	CN759569	ID0AAA25D	
c 704	14	70.0	735	8	DR688913	c 776	14	70.0	802	7	CN759569	ID0AAA25D	
c 705	14	70.0	737	8	BQ837963	c 777	14	70.0	803	6	CB994865	AGENCOURT	
c 706	14	70.0	737	8	DN041590	c 778	14	70.0	804	8	CX325732	JGI_XZT14	
c 707	14	70.0	740	7	CF867565	c 779	14	70.0	804	8	CX325732	JGI_XZT14	
c 708	14	70.0	743	7	CR443180	c 780	14	70.0	806	6	CB631069	OS1EB08G	
c 709	14	70.0	743	10	CL797422	c 781	14	70.0	808	6	CB994647	AGENCOURT	
710	14	70.0	744	8	CX484349	c 782	14	70.0	812	7	CN065247	C09_Ag2_P	
711	14	70.0	746	7	CR281948	c 783	14	70.0	812	7	CN065247	C09_Ag2_P	
712	14	70.0	748	8	DN587115	c 784	14	70.0	813	9	CX927674	JGI_CAAAN1	
713	14	70.0	749	6	CD485795	c 785	14	70.0	813	9	CX927674	JGI_CAAAN1	
714	14	70.0	751	8	DN015937	c 786	14	70.0	816	8	CB914245	JGI_CAAAN4	
c 715	14	70.0	752	6	CA264701	c 787	14	70.0	816	9	CL016993	PUDGY68TD	
c 716	14	70.0	752	10	CL581907	c 788	14	70.0	818	10	CL755415	OR_BBa012	
c 717	14	70.0	753	5	BX779141	c 790	14	70.0	821	7	CV658757	gmrtDnR50	
c 718	14	70.0	756	5	BQ743229	c 791	14	70.0	821	7	CV658757	gmrtDnR50	
c 719	14	70.0	756	9	BZ276796	c 792	14	70.0	822	8	CX897486	JGI_CAAAM7	
c 720	14	70.0	758	5	BW366880	c 793	14	70.0	823	9	BZ978597	PUDGY68TD	
c 721	14	70.0	759	2	BG702633	c 794	14	70.0	824	10	CG195203	PUDGY68TD	
c 722	14	70.0	759	5	BX672760	c 795	14	70.0	825	9	BZ750229	PUDGY68TD	
c 723	14	70.0	760	3	BI560595	c 796	14	70.0	828	6	CA257816	SCCCT300	
c 724	14	70.0	763	7	CK142371	c 797	14	70.0	828	10	CZ280415	cp11c11.r	
c 725	14	70.0	763	8	CX368412	c 798	14	70.0	830	9	CZ280415	cp11c11.r	
c 726	14	70.0	764	6	CB655021	c 799	14	70.0	830	9	CZ280415	cp11c11.r	
c 727	14	70.0	765	8	CX388999	c 800	14	70.0	833	7	CK312184	SB02011A1	
c 728	14	70.0	765	8	CX510994	c 801	14	70.0	833	10	CZ240622	A1AA-aad2	
c 729	14	70.0	767	10	CW440612	c 802	14	70.0	835	9	CK806431	JGI_CAAJ1	
c 730	14	70.0	769	7	CN093572	c 803	14	70.0	835	9	BZ138565	CH230-288	
c 731	14	70.0	770	2	BG297559	c 804	14	70.0	836	2	BE277100	601178605	
c 732	14	70.0	771	10	AG580785	c 805	14	70.0	836	3	BE277100	601178605	
c 733	14	70.0	772	8	CX928802	c 806	14	70.0	840	7	CO115276	GR_EB016	
c 734	14	70.0	772	9	BZ748689	c 807	14	70.0	841	10	CW457855	fbdb001f2	
c 735	14	70.0	772	9	CE668904	c 808	14	70.0	841	10	CW457855	fbdb001f2	
c 736	14	70.0	773	10	CZ134021	c 809	14	70.0	844	7	DN975669	MT08_3_Ta	
c 737	14	70.0	774	7	CJ342363	c 810	14	70.0	845	7	CO452007	CZ08_3_Ta	
c 738	14	70.0	774	8	DR036557	c 811	14	70.0	845	8	CX381930	JGI_XZT53	
c 739	14	70.0	775	9	CX934715	c 812	14	70.0	848	10	CG214797	OG0GC31TH	
c 740	14	70.0	775	10	CL550593	c 813	14	70.0	848	8	DR426783	naw3G06	
c 741	14	70.0	776	10	CL653171	c 814	14	70.0	849	9	CC406035	PULH08TD	
c 742	14	70.0	777	3	BP707048	c 815	14	70.0	851	7	CN093571	EC2BBA8BF	
c 743	14	70.0	778	1	AJ568530	c 816	14	70.0	852	3	BI562686	603256733	
c 744	14	70.0	778	10	CG682797	c 817	14	70.0	852	6	CB995723	AGENCOURT	
c 745	14	70.0	780	5	BQ802996	c 818	14	70.0	852	6	CB995723	AGENCOURT	
c 746	14	70.0	780	5	BZ286048	c 819	14	70.0	853	9	CC335888	OQ0BT55TV	
c 747	14	70.0	786	8	CX921078	c 820	14	70.0	854	10	CG313974	OG3ACB9TH	
c 748	14	70.0	786	9	CC395505	c 821	14	70.0	854	10	CG313974	OG3ACB9TH	
c 749	14	70.0	787	7	CV145775	c 822	14	70.0	854	10	CG313974	OG3ACB9TH	
c 750	14	70.0	787	8	CX806430	c 823	14	70.0	859	8	CX928352	JGI_CAAAN1	
c 751	14	70.0	788	5	BQ788942	c 824	14	70.0	861	8	CV766210	FGA506059	
c 752	14	70.0	788	10	BX157287	c 825	14	70.0	862	5	BU296550	603742202	



C 826	14	70.0	862	8	CX917392	JGI_CAAAG6	C 899	14	70.0	961	10	CL999101	CL999101 ZMMBH6001
C 827	14	70.0	862	10	CW716930	A1FA-aaB7	C 900	14	70.0	962	10	CG135889	CG135889 PUFU45TB
C 828	14	70.0	867	8	DN081756	JGI_CABDI	C 901	14	70.0	964	5	BU411645	BU411645 603153436
C 829	14	70.0	867	9	AZ535579	ENTCF35TF	C 902	14	70.0	966	2	BQ762974	BQ762974 602735029
C 830	14	70.0	867	10	CL503423	SAIL_724	C 903	14	70.0	976	5	BQ713154	BQ713154 AGENCOURT
C 831	14	70.0	868	2	BF541457	602069289-	C 904	14	70.0	980	9	CC443038	CC443038 PUFHC65TB
C 832	14	70.0	869	5	BX750383	BX750383	C 905	14	70.0	987	3	BM050726	BM050726 603633507
C 833	14	70.0	869	8	CW768349	FGAS06274	C 906	14	70.0	987	5	BU512432	BU512432 AGENCOURT
C 834	14	70.0	869	9	BZ810231	PUFPC54TB	C 907	14	70.0	988	2	BB876299	BB876299 601486616
C 835	14	70.0	870	8	CX915585	JGI_CAAAN5	C 908	14	70.0	989	5	BU177551	BU177551 AGENCOURT
C 836	14	70.0	872	10	CG225588	CG225588 OTGGH87TV	C 909	14	70.0	992	5	BQ708400	BQ708400 AGENCOURT
C 837	14	70.0	873	10	AZ683120	ENTG94TF	C 910	14	70.0	992	10	CG165903	CG165903 PUF2X07TD
C 838	14	70.0	873	10	CZ928653	CZ928653 245791 To	C 911	14	70.0	995	5	BQ888754	BQ888754 AGENCOURT
C 839	14	70.0	874	8	CB683462	OSUNEf11N	C 912	14	70.0	1000	6	CD558636	CD558636 AGENCOURT
C 840	14	70.0	874	8	DR566134	WS0266.B2	C 913	14	70.0	1003	9	CC253957	CC253957 CH261-79B
C 841	14	70.0	874	10	AG876930	Oryza sat	C 914	14	70.0	1007	2	BF133515	BF133515 601646052
C 842	14	70.0	875	9	BZ978599	BZ978599 PUGF281TD	C 915	14	70.0	1007	5	BQ712246	BQ712246 AGENCOURT
C 843	14	70.0	876	2	BG394249	602456448	C 916	14	70.0	1016	8	DN940319	DN940319 8678.3 AE
C 844	14	70.0	876	9	BZ044879	lkf54f11.	C 917	14	70.0	1017	9	CG685070	CG685070 OGULH65TV
C 845	14	70.0	877	5	BU374958	603590611	C 918	14	70.0	1018	5	BY703573	BY703573 BY703573
C 846	14	70.0	881	5	BU382210	603855968	C 919	14	70.0	1019	3	BQ049166	BQ049166 AGENCOURT
C 847	14	70.0	883	6	CA814916	CA12E1201	C 920	14	70.0	1023	7	CN645462	CN645462 ILLUMIGN
C 848	14	70.0	883	9	CZ516548	CH240_362	C 921	14	70.0	1024	5	BQ710263	BQ710263 AGENCOURT
C 849	14	70.0	883	10	CZ536806	CZ536806 SRAA-aaD0	C 922	14	70.0	1027	10	CG881937	CG881937 ZMMBB049
C 850	14	70.0	884	6	CD573853	UCRPT01_0	C 923	14	70.0	1030	10	CL116459	CL116459 ISB1-6511
C 851	14	70.0	884	7	CR981226	CR981226	C 924	14	70.0	1061	5	BU200992	BU200992 603104179
C 852	14	70.0	884	8	DR036558	47097.3 L	C 925	14	70.0	1066	2	BG832182	BG832182 602765347
C 853	14	70.0	885	9	CZ516720	CH240_362	C 926	14	70.0	1078	10	CL039353	CL039353 CH216-47M
C 854	14	70.0	885	10	AG877164	Oryza sat	C 927	14	70.0	1080	11	DQ040537	DQ040537 Homo sapi
C 855	14	70.0	890	5	BX327839	BX327839	C 928	14	70.0	1085	3	BM909092	BM909092 AGENCOURT
C 856	14	70.0	890	6	CZ353089	SCJFLA09	C 929	14	70.0	1092	6	CB592554	CB592554 AGENCOURT
C 857	14	70.0	892	6	CD779189	EST650550	C 930	14	70.0	1092	9	CC302910	CC302910 CH261-52H
C 858	14	70.0	892	8	W30144	mc27h09.r1	C 931	14	70.0	1094	5	BQ887568	BQ887568 AGENCOURT
C 859	14	70.0	892	9	CC443041	PUHRC65TD	C 932	14	70.0	1110	6	CF242541	CF242541 AGENCOURT
C 860	14	70.0	893	6	CD245976	CD245976 AGENCOURT	C 933	14	70.0	1112	10	CZ364475	CZ364475 ZMMBF0143
C 861	14	70.0	893	7	CK870361	AGENCOURT	C 934	14	70.0	1116	2	BE906138	BE906138 601497186
C 862	14	70.0	893	9	CG653742	CG653742 OG7AL71TC	C 935	14	70.0	1127	9	CC193298	CC193298 CH261-180
C 863	14	70.0	895	5	BX778015	BX778015 BX778015	C 936	14	70.0	1128	10	CW930402	CW930402 EDC233TF
C 864	14	70.0	895	10	AG159093	AG159093 Pan trogl	C 937	14	70.0	1129	8	DR143641	DR143641 49108859
C 865	14	70.0	896	7	CQ930270	CQ930270 AGENCOURT	C 938	14	70.0	1179	8	DR147106	DR147106 49272374
C 866	14	70.0	896	9	CG694719	CG694719 OGLBQ81TV	C 939	14	70.0	1179	9	CC228614	CC228614 CH261-45J
C 867	14	70.0	898	10	AG873796	Oryza sat	C 940	14	70.0	1201	7	CK209706	CK209706 FGAS02148
C 868	14	70.0	900	7	CO453107	MZCCL1018	C 941	14	70.0	1206	10	AG163926	AG163926 Pan trogl
C 869	14	70.0	900	8	DN081755	DN081755 JGI_CABDI	C 942	14	70.0	1207	5	BU462668	BU462668 603368138
C 870	14	70.0	901	2	BG760411	602716834	C 943	14	70.0	1227	7	CK211045	CK211045 FGAS02287
C 871	14	70.0	901	5	BX327047	BX327047	C 944	14	70.0	1257	5	BU330224	BU330224 603494154
C 872	14	70.0	901	6	CD246686	CD246686 AGENCOURT	C 945	14	70.0	1275	10	CL965865	CL965865 OslFCC012
C 873	14	70.0	911	9	BZ668924	PUBCX89TD	C 946	14	70.0	1280	10	AG435596	AG435596 Mus muscu
C 874	14	70.0	911	10	CG236341	CG236341 OG3DO69TH	C 947	14	70.0	1317	8	DN694146	DN694146 CGX22-EL0
C 875	14	70.0	911	11	CR230253	CR230253 Forward B	C 948	14	70.0	1344	10	CL650080	CL650080 CH213-275
C 876	14	70.0	914	5	BX764118	BX764118	C 949	14	70.0	1359	6	CB308945	CB308945 AGENCOURT
C 877	14	70.0	916	6	CF226157	CF226157 AGENCOURT	C 950	14	70.0	1368	8	CX106125	CX106125 B1125C03
C 878	14	70.0	917	10	CZ512488	CZ512488 GNM2-37L1	C 951	14	70.0	1527	4	AK009500	AK009500 Mus muscu
C 879	14	70.0	918	10	CG920204	CG920204 MBKML0TF	C 952	14	70.0	1565	2	BE421561	BE421561 HWM0115B
C 880	14	70.0	919	6	CF216534	AGENCOURT	C 953	14	70.0	1692	2	BF234458	BF234458 602028253
C 881	14	70.0	921	5	BU318035	BU318035 603951650	C 954	14	70.0	2022	10	AY417450	AY417450 Homo sapi
C 882	14	70.0	924	9	CG637048	CG637048 OGUKV12TV	C 955	14	70.0	2048	7	CN970564	CN970564 18061.123
C 883	14	70.0	924	9	CG995831	CG995831 ZUACGB7TH	C 956	14	70.0	2705	11	DQ030436	DQ030436 Homo sapi
C 884	14	70.0	925	10	CG448340	CG448340 OG5EJ06TV	C 957	14	70.0	3728	4	AK036402	AK036402 Mus muscu
C 885	14	70.0	928	7	CK125211	BE8182410	C 958	14	70.0	3728	4	AK036784	AK036784 Mus muscu
C 886	14	70.0	928	10	CG313989	CG313989 OG3AC89TV	C 959	14	70.0	5194	4	CR859162	CR859162 Pongo pyg
C 887	14	70.0	934	6	CF407156	CF407156 CH3#046 F	C 960	14	70.0	5693	4	CR627061	CR627061 Homo sapi
C 888	14	70.0	934	7	CN063601	CN063601 Ag2_P43 A	C 961	13	65.0	92	9	BH630861	BH630861 1007095C0
C 889	14	70.0	936	10	CZ998510	CZ998510 187882 To	C 962	13	65.0	93	6	CF358169	CF358169 rm87c09.Y
C 890	14	70.0	939	10	CL076393	CL076393 CH216-139	C 963	13	65.0	105	9	AZ519973	AZ519973 RPCI-11-7
C 891	14	70.0	940	10	CZ334544	CZ334544 ZMMBF0043	C 964	13	65.0	105	9	CC407041	CC407041 PUFHQZ93TB
C 892	14	70.0	945	10	CG308521	CG308521 OGZAM31TH	C 965	13	65.0	109	9	BH143820	BH143820 TDGEC28TH
C 893	14	70.0	946	6	CD386447	CD386447 AGENCOURT	C 966	13	65.0	110	1	A1349564	A1349564 qp73c09.x
C 894	14	70.0	949	9	CG685063	CG685063 OGULH55TH	C 967	13	65.0	113	2	BF942085	BF942085 nae86e08.
C 895	14	70.0	952	5	BX411337	BX411337 BX411337	C 968	13	65.0	130	3	BJ749105	BJ749105 BJ749105
C 896	14	70.0	954	10	CZ693499	CZ693499 L008 SP23	C 969	13	65.0	140	7	CO855515	CO855515 LM SL5_00
C 897	14	70.0	955	10	CG316879	CG316879 OGVR50TV	C 970	13	65.0	141	9	AZ758345	AZ758345 IM0550L13
C 898	14	70.0	958	6	CA275372	CA275372 SCBPSD103	C 971	13	65.0	148	2	BF367767	BF367767 RC2-GN003

```
972 13 65.0 148 6 CF308169 ABF--01-O
c 973 13 65.0 156 9 A034407 CIT-HSP-2
974 13 65.0 156 10 BX291463 Arabidops
c 975 13 65.0 157 6 CB278907 ru30b02.Y
c 976 13 65.0 157 8 DN985850 MSU 28F.2
c 977 13 65.0 159 6 CA910319 PCS01679x
c 978 13 65.0 161 1 AA887300 oJ48908.s
979 13 65.0 161 10 CW346948 f8bb001f0
c 980 13 65.0 163 10 BX998489 Forward s
c 981 13 65.0 165 9 AZ315631 1M0032K24
982 13 65.0 170 2 BF874470 RC1-ET013
c 983 13 65.0 173 2 BE172157 RHO-HT055
c 984 13 65.0 173 3 BI320503 sah56Q05.
985 13 65.0 173 5 BW475846 BW475846
c 986 13 65.0 175 3 BM874193 laa02f08.
c 987 13 65.0 177 1 AV552054 AV552054
c 988 13 65.0 182 7 CN795023 DMP0153 M
c 989 13 65.0 182 10 AG213005 Oryza sat
c 990 13 65.0 191 2 BF460685 UI-M-CG0p
991 13 65.0 194 1 AU304561 AU304561
c 992 13 65.0 199 1 BB183251 BB183251
c 993 13 65.0 200 10 CL274920 Ggal.88b
994 13 65.0 204 2 BE955497 UI-M-BH4-
c 995 13 65.0 204 10 AG205720 Oryza sat
996 13 65.0 205 8 DN251640 ACAB-aab9
c 997 13 65.0 206 5 BY154361 BY154361
c 998 13 65.0 207 5 BQ610872 sap51h05.
c 999 13 65.0 208 10 AG955822 Drosophil
c1000 13 65.0 209 9 A2587012 1M0394B21

ALIGNMENTS
RESULT 1
W35284 191 bp mRNA linear EST 13-MAY-1996
LOCUS W35284.1 GI:1317280
DEFINITION W35284.1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone
(HUMAN); mRNA sequence.
ACCESSION W35284.1
VERSION W35284.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 191)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfsing,T., Soares,M., Tan,F.,
Trevasakis,B., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Seq primer: mob.REGA+ET
High quality sequence stop: 1.
FEATURES
source
1. .191
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="taxon:9606"

CF308169 ABF--01-O
A034407 CIT-HSP-2
BX291463 Arabidops
CB278907 ru30b02.Y
DN985850 MSU 28F.2
CA910319 PCS01679x
AA887300 oJ48908.s
CW346948 f8bb001f0
BX998489 Forward s
AZ315631 1M0032K24
BF874470 RC1-ET013
BE172157 RHO-HT055
BI320503 sah56Q05.
BW475846 BW475846
BM874193 laa02f08.
AV552054 AV552054
CN795023 DMP0153 M
AG213005 Oryza sat
BF460685 UI-M-CG0p
AU304561 AU304561
BB183251 BB183251
CL274920 Ggal.88b
BE955497 UI-M-BH4-
AG205720 Oryza sat
DN251640 ACAB-aab9
BY154361 BY154361
BQ610872 sap51h05.
AG955822 Drosophil
A2587012 1M0394B21

/clone="IMAGE:321607"
/tissue_type="parathyroid tumor"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares parathyroid tumor NbHPA"
/notes="Organ: parathyroid gland; Vector: pT7T3D
(Pharmacia) with a modified polylinker; Site1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer
[5'-TGTTACCAATCTGAGTGGGAGCGCGCACCAATTTTTTTTTTTTTTTT
TTTTT-3'], double-stranded cDNA was size selected, ligated
to Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT7T3
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."
```

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 230;  
 Best Local Similarity 100.0%; Pred. No. 0.3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTGGGCAACACAT 20  
 ||||||||||||||||||  
 Db 139 GCTTTGGTGGGCAACACAT 158

## RESULT 3

T85080 231 bp mRNA linear EST 17-MAR-1995  
 LOCUS ye03g09.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
 DEFINITION IMAGE:116704.3' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3  
 (HUMAN); mRNA sequence.

## ACCESSION

VERSION T85080.1 GI:713432

## KEYWORDS

EST.

## SOURCE

ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

## REFERENCE

AUTHORS 1 (bases 1 to 231)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
 Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and  
 Wilson, R.

## TITLE

The WashU-Merck EST Project

## JOURNAL

Unpublished (1995)

## COMMENT

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

Insert Size: 914

High quality sequence stops: 33 Source: IMAGE Consortium, LLNL This  
 clone is available royalty-free through LLNL; contact the IMAGE  
 Consortium (info@image.llnl.gov) for further information.

Insert Length: 914 Std Error: 0.00

Seq primer: -21ml3

High quality sequence stop: 33.

## FEATURES

source

1..231  
 /location/Qualifiers  
 1..231  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:472321"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:116704"  
 /sex="male"

/dev\_stage="20 week-post conception fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares fetal liver spleen INFLS"  
 /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)  
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
 1st strand cDNA was primed with a Pac I - oligo(dT) primer  
 [5' AACTGGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Pac I and cloned into the Pac I  
 and Eco RI sites of the modified pT7T3 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M.Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 8; Length 231;  
 Best Local Similarity 100.0%; Pred. No. 0.3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTGGGCAACACAT 20

Db 167 GCTTTGGTGGGCAACACAT 186  
 ||||||||||||||||||

## RESULT 4

AU023800 238 bp mRNA linear EST 20-OCT-1998  
 LOCUS AU023800 Mouse unfertilized egg cDNA Mus musculus cDNA clone  
 DEFINITION J0439B03.3', mRNA sequence.

## ACCESSION

VERSION AU023800.1 GI:3394147

## KEYWORDS

EST.

## SOURCE

Mus musculus (house mouse)

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

AUTHORS 1 (bases 1 to 238)  
 Ko, M.S.H, Kitchen, J.R., Wang, X., Threat, T.A., Sun, T., DePalma, G.E.,  
 Liang, Y., Kargul, G.J., Sharara, R. and Doi, H.  
 Systematic analyses of genes expressed in unfertilized mouse eggs  
 (The ERATO/Doi Project at Wayne State University) (Ko, M.S.H. et  
 al.)

## JOURNAL

Unpublished (1998)

## COMMENT

Contact: Hirofumi Doi

Doi Biosymmetry Project, ERATO

Japan Science and Technology Corporation (JST)

WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan

Email: hdbioa.jst.go.jp.

## FEATURES

source

1..238  
 /location/Qualifiers  
 1..238  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="J0439B03"  
 /dev\_stage="unfertilized egg"  
 /clone\_lib="Mouse unfertilized egg cDNA"

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 0.3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTGGGCAACACAT 20

||||||||||||||||||

Db 156 GCTTTGGTGGGCAACACAT 175

## RESULT 5

H82288 240 bp mRNA linear EST 09-NOV-1995  
 LOCUS YV79h07.s1 Soares melanocyte 2Nbhm Homo sapiens cDNA clone  
 DEFINITION IMAGE:248989.3' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3  
 (HUMAN); mRNA sequence.

## ACCESSION

VERSION H82288.1 GI:1060377

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

## REFERENCE

AUTHORS 1 (bases 1 to 240)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
 Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and  
 Wilson, R.

## TITLE

The WashU-Merck EST Project

## JOURNAL

Unpublished (1995)

## COMMENT

Contact: Wilson RK  
 Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 1

High quality sequence stops: 1

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Seq primer: Promega -21ml3

High quality sequence stop: 1.

Location/Qualifiers

1. .240

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:3866695"

/db\_xref="taxon:9606"

/clone="IMAGE:249899"

/sex="Male"

/tissue type="melanocyte"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares melanocyte 2NBRM"

/note=Vector: pT73D (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGGAGCGCGAGTCTTTTCTTTTCTTTT 3'],

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT73 vector

(Pharmacia). Library constructed by Bento Soares and

M.Patino Bonaldo. RNA from normal foreskin melanocytes

(FS374) was kindly provided by Dr. Anthony P. Albino."

## ORIGIN

Query Match 100.0%; Score 20; DB 8; Length 240;

Best Local Similarity 100.0%; Pred. No. 0.3;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTGGTGGGCAACACAT 20

Db 161 GCTTGGTGGGCAACACAT 180

## RESULT 6

BB115146/c

LOCUS

DEFINITION BB115146 RIKEN full-length enriched, adult male urinary bladder Mus

musculus cDNA clone 9530049D21 3' similar to AF114258 Mus musculus

forkhead protein FKHR1 (FKHR1) mRNA, mRNA sequence.

BB115146

BB115146.1 GI:8767714

EST.

Mus musculus (house mouse)

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Muridae; Murinae; Mus.

1 (bases 1 to 255)

Kanno.H., Aizawa.K., Akahira.S., Akiyama.J., Arakawa.T.,

Carninci.P., Endo.T., Fukuda.S., Fukunishi.Y., Hara.A., Hayatsu.N.,

Hirozane.T., Hori.F., Ishii.Y., Ishikawa.J., Ishikawa.T., Itoh.M.,

Izawa.M., Kadota.K., Kagawa.I., Kai.C., Kawai.J., Kikuchi.N.,

Kiyosawa.H., Kojima.Y., Kondo.S., Koya.S., Kurihara.C.,

Kusakabe.M., Matsuyama.T., Miki.R., Mizuno.Y., Nakamura.M., Oda.H.,

Okazaki.Y., Ono.T., Owa.C., Saito.H., Sakai.C., Sato.K.,

Shibata.K., Shibata.Y., Shigemoto.Y., Shinagawa.A., Shiraki.T.,

Soqabe.Y., Sugahara.Y., Suzuki.H., Suzuki.H., Tagawa.A.,

Takahashi.F., Tomimaga.N., Toya.T., Tsunoda.Y., Watanishi.A.,

Watanabe.S., Yamamura.T., Yamanaka.I., Yano.R., Yasunishi.A.,

Yokota.T., Yoshida.K., Yoshiaki.A., Yoshino.M., Muramatsu.M. and

Hayashizaki.Y.

RIKEN Mouse ESTs (Konno.H., et al.)

## TITLE

## JOURNAL

## COMMENT

Unpublished (2000)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoaka,S.,

Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki.Y.

Thermolabile and thermoactivation of thermolabile enzymes by

trehalose and its application for the synthesis of full length

cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh,M., Kitsuina,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,

Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,

Okazaki,Y. and Hayashizaki,Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for

further details.

## FEATURES

### source

1. .255

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="9530049D21"

/sex="male"

/tissue type="urinary bladder"

/dev stage="adult"

/lab\_host="DH10B"

/clone\_lib="RIKEN full-length enriched, adult male urinary

bladder"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGAGAGAAGCATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was

prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by

cap-trapper. cDNA went through one round of normalization

to Rot = 20.0 and subtraction to Rot = 370.0. Second

strand cDNA was prepared with the primer adapter of

sequence [5' GAGAGAGATTCTCGAGTTAATAATATCCCTCCCTCC

3'] cDNA was cleaved with XhoI and BamHI. Vector: a

modified pBluescript KS(+) after bulk excision from

Lambda FLC I."

## ORIGIN

Query Match

Best Local Similarity 100.0%; Score 20; DB 1; Length 255;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTGGTGGGCAACACAT 20

Db 100 GCTTGGTGGGCAACACAT 81

## RESULT 7

AV318369/c

LOCUS

DEFINITION AV318369 RIKEN full-length enriched, 13 days embryo

musculus cDNA clone 5930401I20 3' similar to AF114258 Mus musculus

forkhead protein FKHR1 (FKHR1) mRNA, mRNA sequence.

AV318369

ACCSSION

VERSION AV318369.1 GI:6283786

KEYWORDS EST.



/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTAATTAATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda F1C I. Cloning sites, 5' end: Sali; 3' end: BamHI."

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 287;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTGGTGGCAACACAT 20  
|||||  
Db 132 GCTTGGTGGCAACACAT 113

## RESULT 9

CF135310  
LOCUS  
DEFINITION UI-HP-BROp-ary-a-08-0-UI.r1 NIH\_MGC\_52 Homo sapiens cDNA clone IMAGE:30554023 5', mRNA sequence.

ACCESSION CF135310  
VERSION CF135310.1 GI:33250754  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 314)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)  
PUBMED 889548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565

Email: bento-soares@uiowa.edu  
Tissue Procurement: Louis Staudt  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/humanfl.html>  
Seq primer: pYX-5.

## FEATURES

source  
Location/Qualifiers  
1..314  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30554023"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/cell\_line="MGC85"  
/lab\_host="DH10B (LTI)"  
/clone\_lib="NIH\_MGC\_52"

/note="Vector: p7T3-Pac; Site 1: NotI; Site 2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA (7.4-9.5kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D. "

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 314;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTGGTGGCAACACAT 20  
|||||  
Db 168 GCTTGGTGGCAACACAT 187

## RESULT 10

BU783931/c  
LOCUS  
DEFINITION BU783931 in11901.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:6124224 3', similar to SW:FKHR\_HUMAN Q12778 FORK HEAD DOMAIN PROTEIN FKHR. [1]; mRNA sequence.

ACCESSION BU783931  
VERSION BU783931.1 GI:23828583  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 316)

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Brestelli,J., Gradwohl,J., Clifton,S., Hillier,L., Marra,M., Page,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,S., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T., Jackson,Y. and Bowers,Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)

TITLE  
JOURNAL  
COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from Gibco.

Location/Qualifiers

1..316

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6124224"

/tissue\_type="insulinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="Human insulinoma"

/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:

XhoI; Site 2: EcoRI; Constructed with lambda ZapII system

(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to

pBluescript SK- by Dr. H. Inoue following the Washington

University protocol

([http://genome.wustl.edu/est/lambda\\_protocol.shtml](http://genome.wustl.edu/est/lambda_protocol.shtml)).

Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt

Laboratory, Washington University School of Medicine, Box

8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this

is a Washington University Pancreas EST project library. "



SOURCE  
ORGANISM Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS 1 (bases 1 to 349)  
Henrich, J., Hermanns, J., Kranz, H., Loebbert, R., Schlueter, T.,  
Schuette, D., Weindel, M., Heil, O., Ebert, L., Neubert, P., Peters, M.,  
Radelof, U., Schneider, D. and Korn, B.  
TITLE Mouse ArrayTAG cDNA (LION)  
JOURNAL Unpublished (2003)  
COMMENT Contact: Ina Kolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; LIONp462B0882.  
RZPDLIB;  
Mouse ArrayTAG cDNA (LION)  
http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi?response?libNo=4  
62 Contact: Ina Kolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de  
This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
RP: CAGGAACAGCTATGAC.  
FEATURES  
source  
Location/Qualifiers  
1..349  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="LIONp462B0882"  
/lab\_host="DH10B"  
/clone\_lib="pbluescript Lion"  
ORIGIN  
Query Match 100.0%; Score 20; DB 5; Length 349;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 161 GCTTTGGTTGGCAACACAT 180  
RESULT 14  
W32908/c  
LOCUS W32908 353 bp mRNA linear EST 13-MAY-1996  
DEFINITION zc07d04.x1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone  
IMAGE:321607 5' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3  
(HUMAN); mRNA sequence.  
ACCESSION W32908.1 GI:1314963  
VERSION W32908.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 353)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, P.,  
Trevaaskis, B., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu  
This clone is available royalty-free through LILN; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: mob.REGA+ET  
High quality sequence stop: 328.  
Location/Qualifiers  
1..353  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:1259265"  
/db\_xref="taxon:9606"  
/clone="IMAGE:321607"  
/tissue type="parathyroid tumor"  
/dev stage="adult"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares parathyroid tumor NbHPA"  
/note="Organ: parathyroid gland; Vector: pRTT3D  
(Pharmacia) with a modified polylinker; Site 1: Not I;  
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -  
oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGGCGCGCCACCAATTTTTTTTTTTTTTTT  
TTTGT-3'] double-stranded cDNA was size selected, ligated  
to Eco RI adapters (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of a modified pRTT3  
vector (Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid  
adenomas was kindly provided by Dr. Stephen Marx, National  
Institute of Diabetes and Digestive and Kidney Diseases,  
NIH."  
ORIGIN  
Query Match 100.0%; Score 20; DB 8; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 180 GCTTTGGTTGGCAACACAT 161  
RESULT 15  
CR469238 355 bp mRNA linear EST 01-JUL-2004  
LOCUS CR469238 Rat pBluescript Lion Rattus norvegicus cDNA clone  
DEFINITION LIONp463E08441 3', mRNA sequence.  
ACCESSION CR469238  
VERSION CR469238.1 GI:49601587  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 355)  
Henrich, J., Hermanns, J., Kranz, H., Loebbert, R., Schlueter, T.,  
Schuette, D., Weindel, M., Heil, O., Ebert, L., Neubert, P., Peters, M.,  
Radelof, U., Schneider, D. and Korn, B.  
TITLE Rat ArrayTAG cDNA  
JOURNAL Unpublished (2004)  
COMMENT Contact: Inge Arlart  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Email: www.rzpd.de  
RZPD; LIONp463E08441.  
RZPDLIB;  
Rat ArrayTAG cDNA  
http://www.rzpd.de/cgi-  
bin/products/showLib.pl.cgi?response?libNo=463 Contact: Inge Arlart  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 100  
Fax: +49 30 32639 111



```

www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
RP: CAGGAACAGCTATGAC.
FEATURES
    source
        Location/Qualifiers
            1..355
                /organism="Rattus norvegicus"
                /mol_type="mRNA"
                /db_xref="taxon:10116"
                /clone="U010P463E08441"
                /lab_host="DH10B"
                /clone_lib="Rat pBluescript Lion"
ORIGIN
Query Match      100.0%; Score 20; DB 7; Length 355;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAACACAT 20
|||||
Db 159 GCTTTGGTTGGGCAACACAT 178

RESULT 16
H99327
LOCUS
DEFINITION
Yx23c01.s1 Soares melanocyte 2N6HM Homo sapiens cDNA clone
IMAGE:262560 3' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3
(HUMAN); mRNA sequence.
ACCESSION
H99327
VERSION
H99327.1 GI:1123995
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 358)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 233
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 2771 Std Error: 0.00
Seq primer: ml3 -40 forward
High quality sequence stop: 233.
FEATURES
    source
        Location/Qualifiers
            1..358
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="GDB:3872202"
                /db_xref="taxon:9606"
                /clone="IMAGE:262560"
                /sex="Male"
                /tissue_type="melanocyte"
                /lab_host="DH10B (ampicillin resistant)"
                /clone_lib="Soares melanocyte 2N6HM"
                /note="Vector: pTT3D (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCGCGAGTTTTTTTTTTTTTTT 3'],

```

```

double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pTT3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo. RNA from normal foreskin melanocytes
(FS374) was kindly provided by Dr. Anthony P. Albino."
Query Match      100.0%; Score 20; DB 8; Length 358;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAACACAT 20
|||||
Db 163 GCTTTGGTTGGGCAACACAT 182

RESULT 17
H99327
LOCUS
DEFINITION
UI-M-AN1-afi-h-10-0-UI.s1 NIH BMAP MBG N Mus musculus cDNA clone
UI-M-AN1-afi-h-10-0-UI 3', mRNA sequence.
ACCESSION
H99327
VERSION
H99327.1 GI:5490627
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 361)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
normalized basal ganglia library cDNA Library Preparation: M.B.
Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
available by the means for distribution of the BMAP cDNA clones, this
determines the means for distribution of the BMAP cDNA clones, this
record will be updated accordingly when that means is determined.
The following repetitive elements were found in this cDNA sequence:
1-21, >AT-rich#Low complexity
Seq primer: M13 Forward
POLYA=Yes.
FEATURES
    source
        Location/Qualifiers
            1..361
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UI-M-AN1-afi-h-10-0-UI"
                /dev_stage="27-32 days"
                /lab_host="DH10B (Life Technologies)"
                /clone_lib="NIH BMAP MBG N"
                /note="Vector: pTT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH BMAP MBG N library is a normalized library constructed
from mouse basal ganglia. The tag is a string of 5
nucleotides present between the Not I site and the
oligo-dT track. The library was constructed as described

```

by Bonaldo, Lennon and Soares, Genome Research 6:  
791-806, 1996. Tissue provided by Ms. Annie Novakovich,  
Zivic-Miller Laboratories.  
TAG TISSUE=basal-ganglia  
TAG\_LIB=NIH BMAP\_MBG\_N  
TAG\_SEQ=GTGAC

ORIGIN  
Query Match 100.0%; Score 20; DB 1; Length 361;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 173 GCTTTGGTTGGCAACACAT 192  
|||||

RESULT 18  
AA035406  
LOCUS  
DEFINITION  
zk26h09.s1 Soares pregnant uterus NBHPU Homo sapiens cDNA clone  
IMAGE:471713 3' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3  
(HUMAN);, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AA035406.1 GI:1507063  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS  
1 (bases 1 to 382)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevaaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)

TITLE  
JOURNAL  
PUBMED  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 902 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 192.

FEATURES  
source  
1. .382  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3757715"  
/db\_xref="taxon:9606"  
/clone="IMAGE:471713"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="Soares\_pregnant uterus NBHPU"

/note="Organ: uterus; Vector: pT7T3-Pac; Site:1: Not 1;  
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -  
oligo(dT) primer [5,  
AACTGGAGAAATTCGCGCCCTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization. Library  
constructed by M. Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 382;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 159 GCTTTGGTTGGCAACACAT 178  
|||||

## RESULT 19

AW274010  
LOCUS  
DEFINITION  
xv26c05.x1 Soares NFL T GBC\_S1 Homo sapiens cDNA clone  
IMAGE:2814248 3' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3  
(HUMAN);, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AW274010.1 GI:6661040  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 388)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40UP from Gibco.

FEATURES  
source  
1. .388  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2814248"  
/lab\_host="DH10B"  
/clone\_lib="Soares NFL T GBC\_S1"  
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
Equal amounts of plasmid DNA from three normalized  
libraries (fetal lung NBHL19W, testis NHT, and B-cell  
NCI-CGAP GCBI) were mixed, and ss circles were made in  
vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of  
I.M.A.G.E. clones 297480-302087, 582632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 388;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 162 GCTTTGGTTGGCAACACAT 181  
|||||

## RESULT 20

AA981734  
LOCUS  
DEFINITION  
ua30g09.r1 Soares mammary gland NbMMG Mus musculus cDNA clone  
IMAGE:1348288 5' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3  
(HUMAN);, mRNA sequence.

ACCESSION  
AA981734

```

VERSION AA981734.1 GI:3160503
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNLML; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:697080
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 173.
Location/Qualifiers
1..393
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1348288"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NDMMG"
/note="Organ: mammary gland; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTCACCATCTGAAGTGGAGCGCGCGAATGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

ORIGIN
Query Match 100.0%; Score 20; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20
|||||
DB 158 GCTTTGGTTGGCAACACAT 177

RESULT 21
AK212379/c
LOCUS AK212379
DEFINITION Mus musculus cDNA, clone:Y2G0124M02, strand.minus.
reference:ENSEMBL:Mouse-Transcript-ENST:ENSMUST0000053764, based
on BLAT search.
ACCESSION AK212379
VERSION AK212379.1 GI:56036556
KEYWORDS HTC; ASSETS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Watahiki,A., Waki,K., Hayatsu,N., Shiraki,T., Kondo,S.,
Nakamura,M., Sasaki,D., Arakawa,T., Kawai,J., Harbers,M.,
Hayashizaki,Y. and Carninci,P.
TITLE Libraries enriched for alternatively spliced exons reveal splicing
patterns in melanocytes and melanomas
JOURNAL Nat. Methods 1, 233-239 (2004)
REFERENCE
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Harbers,M., Hayatsu,N.,
Hori,F., Imotani,K., Kawai,J., Kondo,S., Murata,M., Nakamura,M.,
Nomura,K., Ohno,M., Sasaki,D., Shiraki,T., Waki,K., Watahiki,A. and
Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-2004) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT Alternative Splicing Libraries (ASLs) are prepared by: Preparing
of single-stranded DNA using a RNA template from full length cDNA
of libraries, hybridizing of
single-stranded DNAs, removing of remaining single-stranded DNA,
digesting of regions comprising double-stranded DNA by a set of 4
bp-cutters, capturing of DNA hybrids with loop structures
(alternative spliced exon), ligating of Y-shaped primers to
isolated DNA hybrids with loop structures, PCR amplification of
ligation products and their cloning into pFLCI vector. (Reference).
Location/Qualifiers
1..396
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="Y2G0124M02"
/cell_lines="mixture of B16-F10Y and melan-c"
/cell_type="mixture of melanoma cell and melanocyte cell"
/clone_lib="Alternative Splicing Library L3"
/note="strand.minus, reference:ENSEMBL:Mouse-Transcript-
ENST:ENSMUST0000053764, based on BLAT search"

FEATURES
source
1..396
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="Y2G0124M02"
/cell_lines="mixture of B16-F10Y and melan-c"
/cell_type="mixture of melanoma cell and melanocyte cell"
/note="strand.minus, reference:ENSEMBL:Mouse-Transcript-
ENST:ENSMUST0000053764, based on BLAT search"

ORIGIN
Query Match 100.0%; Score 20; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20
|||||
DB 230 GCTTTGGTTGGCAACACAT 211

RESULT 22
CF356230/c
LOCUS CF356230
DEFINITION CF356230
maJ53all.y1 McCarrey Eddy 6 day primitive type A spermatogonia Mus
musculus cDNA clone IMAGE:6925620 5' similar to TR:Q9WVH5 Q9WVH5
FORKHEAD PROTEIN FKHL1. [1] ; mRNA sequence.
ACCESSION CF356230
VERSION CF356230.1 GI:34023168
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 396)
McCarrey,J., Eddy,M., Marra,M., Hillier,L., Clifton,S., Pape,D.,
Martin,J., Wylie,T., Dante,M., Bowers,Y., Theising,B., Gibbons,M.,
Ritter,E., Tsagareishvili,R., Ronko,I., Maguire,L., Kennedy,S.,
Bennett,J., Waterston,R. and Wilson,R.
TITLE NIEHS Mouse

```

JOURNAL  
COMMENT

Unpublished (2002)  
Contact: McCarrey/Eddy NIEHS Mouse  
NIEHS Mouse  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Library constructed and donated by J. McCarrey, Ph.D. (Southwest  
Foundation for Biomedical Research, Dept. of Genetics) - excision  
done by E.M. Eddy, Ph.D. (National Institutes of Health, National  
Institute of Environmental Health Sciences).  
Seq primer: Primer name ambiguous.

## FEATURES

source

```
1. .396
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clones="IMAGE:6925620"
/sex="male"
/tissue_type="6-day primitive type A spermatogonia"
/lab_host="DH10B (phage-resistant)"
/clone_lib="McCarrey Eddy 6 day primitive type A
spermatogonia"
/note="Organ: testis; Vector: pBluescript SK+
(Stratagene); Site 1: EORI; Site 2: XhoII; cDNA oligo
dt-primed [5'-(GA)10-ACTAGTCGAGTTTCTTTT-3'] and
directionally cloned using 5' linkers 5'-AATTGGCAGAG-3'
and 5'-CTCGTGGC-3'. Size selection of >400bp material
gives excised insert size ranging from 1-2 Kb. Library was
single-stranded phagemids were prepped and transformed into
DH10B. Library constructed and donated by J. McCarrey,
Ph.D. (Southwest Foundation for Biomedical Research, Dept.
of Genetics); excision done by E.M. Eddy, Ph.D. (National
Institutes of Health, National Institute of Environmental
Health Sciences)."
```

## ORIGIN

```
Query Match      100.0%; Score 20; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

## Qy

```
1 GCTTTGGTTGGCAACACAT 20
```

## Db

```
241 GCTTTGGTTGGCAACACAT 222
```

## RESULT 23

BP755053

## LOCUS

```
DEFINITION BP755053 mouse (C57BL/6) pancreatic islet library with
recombination-based method Mus musculus cDNA clone mial9094 3',
mRNA sequence.
```

## ACCESSION

BP755053

## VERSION

BP755053.1

## KEYWORDS

EST.

## SOURCE

Mus musculus

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Murinae; Mus.

1 (bases 1 to 397)

Nishimura, M., Yokoi, N., Miki, T., Horikawa, Y., Yoshioka, H.,

Takeda, J., Ohara, O. and Seino, S.

Construction of a multi-functional cDNA library specific for mouse

pancreatic islets and its application to microarray

DNA Res. 11 (5), 315-323 (2004)

15747579

Contact: Susumu Seino

Division of Cellular and Molecular Medicine

Kobe University Graduate School of Medicine

7-5-1 Kusunoki-cho, Chuo-ku, Kobe, Hyogo 650-0017, Japan

Tel: 81-78-382-5360

Fax: 81-78-382-5370

Email: seino@med.kobe-u.ac.jp.

## FEATURES

source

```
Location/Qualifiers
1. .397
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="C57BL/6"
/db_xref="taxon:10090"
/clones="mial9094"
/sex="male"
/tissue_type="pancreatic islet"
/dev_stage="adult"
/clone_lib="mouse (C57BL/6) pancreatic islet library with
recombination-based method"
```

## ORIGIN

```
Query Match      100.0%; Score 20; DB 3; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

## Qy

```
1 GCTTTGGTTGGCAACACAT 20
```

## Db

```
156 GCTTTGGTTGGCAACACAT 175
```

## RESULT 24

BY367902/c

## LOCUS

```
DEFINITION BY367902 RIKEN full-length enriched, 6 days neonate spleen Mus
musculus cDNA clone F420013K03 3', mRNA sequence.
```

## ACCESSION

BY367902

## VERSION

BY367902.1

## KEYWORDS

EST.

## SOURCE

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Murinae; Mus.

1 (bases 1 to 397)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,

Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusica, V.,

Chothia, C., Corbani, L. E., Cousins, S., Dalia, E., Dragani, T. A.,

Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Jarvis, E. D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,

Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,

Maltais, K., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J. C., Reid, J., Ring, B. Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tonita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Kasazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Haseizume, W., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,

Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

JOURNAL

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: <http://genome-gsc.riken.jp/>  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H.,  
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, K.,  
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and  
Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site (<http://genome-gsc.riken.go.jp/>) for  
further details.

#### FEATURES

source  
Location/Qualifiers  
1..397  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="F420013K03"  
/tissue\_type="spleen"  
/dev\_stage="6 days neonate"  
/clone\_lib="RIKEN full-length enriched, 6 days neonate  
spleen"

#### ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 397;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAACACAT 20  
|||||  
DB 237 GCTTTGGTTGGGCAACACAT 218

#### RESULT 25

AK211466/c  
LOCUS AK211466 398 bp mRNA linear HTC 23-NOV-2004  
DEFINITION Mus musculus cDNA, clone:Y2G0121N15, strand:plus,  
reference:ENSEMBL:Mouse-Transcript-ENST:ENSMUST0000053764, based  
on BLAT search.

ACCESSION AK211466  
VERSION AK211466.1 GI:56035643

KEYWORDS HTC; ASSETS.

SOURCE Mus musculus (house mouse)

#### ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.

#### REFERENCE

AUTHORS Watahiki, A., Waki, K., Hayatsu, N., Shiraki, T., Kondo, S.,  
Nakamura, M., Sasaki, D., Arakawa, T., Kawai, J., Harbers, M.,  
Hayashizaki, Y. and Carninci, P.

TITLE Libraries enriched for alternatively spliced exons reveal splicing  
patterns in melanocytes and melanomas

JOURNAL Nat. Methods 1, 233-239 (2004)

#### REFERENCE

AUTHORS Arakawa, T., Carninci, P., Fukuda, S., Harbers, M., Hayatsu, N.,  
Hori, F., Imotani, K., Kawai, J., Kondo, S., Murata, M., Nakamura, M.,  
Nomura, K., Ohno, M., Sasaki, D., Shiraki, T., Waki, K., Watahiki, A. and  
Hayashizaki, Y.

#### TITLE

JOURNAL Direct Submission  
Submitted (15-SEP-2004) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa, 230-0045, Japan [E-mail: genome-res@sc.riken.jp,  
URL: <http://genome-gsc.riken.jp>, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216]

#### COMMENT

Alternative Splicing Libraries (ASLs) are prepared by: Preparing  
of single-stranded DNA using a RNA template from full length cDNA  
libraries, hybridizing of  
single-stranded DNAs, removing of remaining single-stranded DNA,  
digesting of regions comprising double-stranded DNA by a set of 4  
bp-cutters, capturing of DNA hybrids with loop structures  
(alternative spliced exon), ligating of Y-shaped primers to  
isolated DNA hybrids with loop structures, PCR amplification of  
ligation products and their cloning into pFLCI vector. (Reference).

#### FEATURES

source

Location/Qualifiers  
1..398  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="Y2G0121N15"  
/cell\_lines="mixture of B16-F10Y and melan-c"  
/cell\_type="mixture of melanoma cell and melanocyte cell"  
/clone\_lib="Alternative Splicing Library L3"  
/note="strand:plus, reference:ENSEMBL:Mouse-Transcript-  
ENST:ENSMUST0000053764, based on BLAT search"

#### ORIGIN

Query Match 100.0%; Score 20; DB 4; Length 398;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCTTTGGTTGGGCAACACAT 20  
|||||  
DB 261 GCTTTGGTTGGGCAACACAT 242

#### RESULT 26

BI401999  
LOCUS BI401999 403 bp mRNA linear EST 14-AUG-2001  
DEFINITION MI-P-CP0-nvu-d-11-0-UI s1 MI-P-CP0 Sus scrofa cDNA clone  
MI-P-CP0-nvu-d-11-0-UI 3', mRNA sequence.

ACCESSION BI401999

VERSION BI401999.1 GI:15181060

KEYWORDS EST.

SOURCE Sus scrofa (pig)

#### ORGANISM

Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;  
Sus.

#### REFERENCE

AUTHORS 1 (bases 1 to 403)  
Bonaldo, M. P., Lennon, G. and Soares, M. B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
PUBMED 8889548

#### COMMENT

Contact: Tuggle CK  
Molecular Genetics Laboratory, Department of Animal Science  
Iowa State University  
201 Klidde Hall, Ames, IA 50011-3150, USA  
Tel: 5152944252  
Fax: 5152942401  
Email: [ckugle@iastate.edu](mailto:ckugle@iastate.edu)

The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the non-normalized uterus library cDNA library preparation: M.B. Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab, University of Iowa. Clone distribution: clones will be available through Research Genetics (www.regen.com) The following repetitive elements were found in this cDNA sequence: 1-21, >AT rich#Low complexity  
Seq primer: M13 Forward  
POLYA=Yes.

#### FEATURES

Location/Qualifiers  
1..403  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/clones="MI-P-CP0-nvu-d-11-0-UI"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="MI-P-CP0"  
/note="Vector: pT7D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: EcoRI; The MI-P-CP0 library is derived from uterus. For a detailed description of the library from which this clone was derived, please visit our web site at <http://pigest.genome.iastate.edu/>. The procedure used to create this library has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG TISSUE=uterus  
TAG LIB=MI-P-CP0  
TAG\_SEQ=AGTCCATCG"

#### ORIGIN

Query Match 100.0%; Score 20; DB 3; Length 403;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 178 GCTTTGGTTGGCAACACAT 197

#### RESULT 27

BB700991/c  
LOCUS BB700991 RIKEN full-length enriched, in vitro fertilized eggs Mus  
DEFINITION musculus cDNA clone 7420426B07 3', mRNA sequence.  
ACCESSION BB700991  
VERSION BB700991.1 GI:16049815  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

#### REFERENCE

AUTHORS Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

#### TITLE

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

#### JOURNAL

COMMENT Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL:<http://genome.gsc.riken.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.  
e mouse tissues.

#### FEATURES

Location/Qualifiers  
1..404  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="7420426B07"  
/sex="Female"  
/tissue type="in vitro fertilized eggs"  
/dev stage="egg"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, in vitro fertilized eggs"  
/note="Site 1: Sall; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'-GAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATCTCGATTATTAATTAATCCCTCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sall; 3' end: BamHI"

#### ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 404;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 249 GCTTTGGTTGGCAACACAT 230

#### RESULT 28

BB835762/c  
LOCUS BB835762 RIKEN full-length enriched, mammary gland RCB-0527  
DEFINITION Jyg-MC(B) cDNA Mus musculus cDNA clone G930038L06 3', mRNA sequence.  
ACCESSION BB835762  
VERSION BB835762.1 GI:17014005  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

**REFERENCE**  
**AUTHORS**  
 Mammalia; Euthera; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
 1 (bases 1 to 411)  
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
 Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K.,  
 Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,  
 Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,  
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,  
 Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
 Tanaka, T., Tomaru, A., Toyota, T., Watahiki, A., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.  
**TITLE**  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)  
**JOURNAL**  
**COMMENT**  
 Unpublished (2001)  
 Contact: Yoshinide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
 and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.  
**FEATURES**  
**source**  
 Location/Qualifiers  
 1..411  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="G930038L06"  
 /tissue\_type="mammary gland"  
 /cell\_line="RCB-0527 Jyg-MC(B)"  
 /clone\_lib="RIKEN full-length enriched, mammary gland  
 RCB-0527 Jyg-MC(B) cDNA"  
**ORIGIN**  
 Query Match 100.0%; Score 20; DB 2; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCTTTGGTTGGCACACAT 20  
 |||||  
 DB 256 GCTTTGGTTGGCACACAT 237  
**REFERENCE**  
**AUTHORS**  
 CF135311 413 bp mRNA linear EST 09-SEP-2003  
 UI-HF-BR0P-ary-a-10-0-UI-ri NIH MGC\_52 Homo sapiens cDNA clone  
 IMAGE:30554025 5', mRNA sequence.  
**TITLE**  
 CF135311  
**JOURNAL**  
**COMMENT**  
 Unpublished (2001)  
 Contact: Yoshinide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
 and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.  
**FEATURES**  
**source**  
 Location/Qualifiers  
 1..411  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="G930038L06"  
 /tissue\_type="mammary gland"  
 /cell\_line="RCB-0527 Jyg-MC(B)"  
 /clone\_lib="RIKEN full-length enriched, mammary gland  
 RCB-0527 Jyg-MC(B) cDNA"  
**ORIGIN**  
 Query Match 100.0%; Score 20; DB 2; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCTTTGGTTGGCACACAT 20  
 |||||  
 DB 256 GCTTTGGTTGGCACACAT 237  
**REFERENCE**  
**AUTHORS**  
 CF135311 413 bp mRNA linear EST 09-SEP-2003  
 UI-HF-BR0P-ary-a-10-0-UI-ri NIH MGC\_52 Homo sapiens cDNA clone  
 IMAGE:30554025 5', mRNA sequence.  
**TITLE**  
 CF135311  
**JOURNAL**  
**COMMENT**  
 Unpublished (2001)  
 Contact: Yoshinide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
 and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.  
**FEATURES**  
**source**  
 Location/Qualifiers  
 1..411  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30554025"  
 /tissue\_type="lymph"  
 /cell\_type="germinal center B cells"  
 /cell\_line="MGC85"  
 /lab\_host="DH10B (LTI)"  
 /clone\_lib="NIH MGC 52"  
 /note="Vector: pVT3-Pac; Site 1: NotI; Site 2: Eco RI;  
 Constructed from size fractionated cytoplasmic mRNA  
 (7.4-9.5kb). Directionally cloned. Cells provided by  
 Louis M. Staudt, Ph.D. Library preparation by Maria de  
 Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

**ORGANISM**  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
**REFERENCE**  
**AUTHORS**  
 Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
**JOURNAL**  
**COMMENT**  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MBRRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Louis Staudt  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/humanfl.html  
 Seq primer: pYX-5.  
**FEATURES**  
**source**  
 Location/Qualifiers  
 1..413  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30554025"  
 /tissue\_type="lymph"  
 /cell\_type="germinal center B cells"  
 /cell\_line="MGC85"  
 /lab\_host="DH10B (LTI)"  
 /clone\_lib="NIH MGC 52"  
 /note="Vector: pVT3-Pac; Site 1: NotI; Site 2: Eco RI;  
 Constructed from size fractionated cytoplasmic mRNA  
 (7.4-9.5kb). Directionally cloned. Cells provided by  
 Louis M. Staudt, Ph.D. Library preparation by Maria de  
 Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

**ORIGIN**  
 Query Match 100.0%; Score 20; DB 6; Length 413;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCTTTGGTTGGCACACAT 20  
 |||||  
 DB 168 GCTTTGGTTGGCACACAT 187  
**RESULT 30**  
 AA692950/c  
**LOCUS**  
 DEFINITION  
 AA692950.1 GI:2693888  
 EST.  
 Mus musculus (house mouse)  
**ACCESSION**  
 VERSION  
 AA692950.1  
**KEYWORDS**  
 SOURCE  
**ORGANISM**  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
**REFERENCE**  
**AUTHORS**  
 1 (bases 1 to 414)  
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
**TITLE**  
 The WashU-HMNI Mouse EST Project  
**JOURNAL**  
 Unpublished (1996)

```
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicinep
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:614340
Putative full length read
vector to vector length is 419
High quality sequence stop: 378.

FEATURES
source
1..414
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:1125004"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/clone_lib="Knowles Solter mouse 2 cell"
/notes="Organ: embryo; Vector: pBluescribe (modified);
Site 1: MluI; Site 2: SalI; Cloned unidirectionally from
mRNA prepared from 13,500 2-cell stage embryos. Primer:
SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTT-3'. cDNAs
were cloned into the MluI/SalI sites of a modified
pBluescribe vector using commercial linkers (NEB).
Average insert size: 1.2 kb."

ORIGIN
Query Match 100.0%; Score 20; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20
|||||
Db 262 GCTTTGGTTGGCAACACAT 243

RESULT 31
AJ647314/c 414 bp mRNA linear EST 07-JUL-2004
LOCUS
DEFINITION
AJ647314 CSEQRAN19 Sus scrofa cDNA clone C0003105_F08, mRNA
sequence.
ACCESSION
AJ647314 GI:49324159
VERSION
AJ647314.1
KEYWORDS
EST.
SOURCE
Sus scrofa (pig)
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.
1 (bases 1 to 414)
Anderson,S.I., Finlayson,H.A. and Archibald,A.L.
Development of cDNA and EST resources for studying reproduction and
embryo development in pigs and cattle
Unpublished (2004)
Contact: Anderson SI
Genomics and Bioinformatics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred
v0.020425.c. Vector identified by cross match with the -minscore 20
and -mismatch 12 options. Vector:pBluescriptII(KS) R. Site1: EcoRI
R. Site2: NotI 5' Seq Primer M13F Normalised library constructed
from pooled ovaries. Clones available from UK Centre for Functional
Genomics in Farm Animals, Roslin Institute, Roslin, Midlothian, UK,
EH25 9PS, www.ark-genomics.org.
Location/Qualifiers
1..414

FEATURES
source
1..414
```

```
ORIGIN
Query Match 100.0%; Score 20; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20
|||||
Db 373 GCTTTGGTTGGCAACACAT 354

RESULT 32
AW825094 418 bp mRNA linear EST 17-MAY-2000
LOCUS
DEFINITION
AW825094 y1 Soares NMGC B-cell Mus musculus cDNA clone
IMAGE:3166712 5' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3
(HUMAN); mRNA sequence.
ACCESSION
AW825094
VERSION
AW825094.1 GI:7918171
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 418)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MGI:1062172
Seq primer: -40RP from Gibco
High quality sequence stop: 393.
Location/Qualifiers
1..418
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3166712"
/tissue_type="germinal B-cell from resting spleen"
/lab_host="DH10B (phage resistant)"
/clone_lib="Soares NMGC B-cell"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCAGGATTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

ORIGIN
Query Match 100.0%; Score 20; DB 1; Length 418;
```



```

Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20
   |||||
Db 101 GCTTTGGTTGGCAACACAT 120

RESULT 33
BG230242 423 bp mRNA linear EST 08-FEB-2001
IMAGE:3384926 3' similar to TR:Q9WVH5 Q9WVH5 FORKHEAD PROTEIN
FKHR1. [1] ; mRNA sequence.
ACCESSION BG230242
VERSION BG230242.1 GI:12717746
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
1 (bases 1 to 423)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1089922
High quality sequence stop: 292.
Location/Qualifiers
1..423
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3384926"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares mouse NMGB bcell"
/clone_lib="Soares mouse NMGB bcell"
/note="Organ: germinal B-cell; Vector: pMT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCGCTGTTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pMT73 vector.
Library is normalized; constructed by Bento Soares and
M.Fatima Bonaudo."
ORIGIN
Query Match 100.0%; Score 20; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20
   |||||
Db 169 GCTTTGGTTGGCAACACAT 188

RESULT 34
BB828031/c
LOCUS BB828031
DEFINITION BB828031 RIKEN full-length enriched, mammary gland RCB-0526
JYG-MC(A) cDNA Mus musculus cDNA clone G830050G14 3', mRNA
sequence.
ACCESSION BB828031
VERSION BB828031.1 GI:17006274
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

425 bp mRNA linear EST 19-NOV-2001
BB828031
LOCUS BB828031
DEFINITION BB828031 RIKEN full-length enriched, mammary gland RCB-0526
JYG-MC(A) cDNA Mus musculus cDNA clone G830050G14 3', mRNA
sequence.
ACCESSION BB828031
VERSION BB828031.1 GI:17006274
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
1 (bases 1 to 425)
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-rs@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
FEATURES
source
1..425
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G830050G14"
/tissue_type="mammary gland"
/cell_line="RCB-0526 Jyg-MC(A)"
/clone_lib="RIKEN full-length enriched, mammary gland
RCB-0526 Jyg-MC(A) cDNA"
ORIGIN
Query Match 100.0%; Score 20; DB 2; Length 425;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20
   |||||
Db 268 GCTTTGGTTGGCAACACAT 249

RESULT 35
CF138380/c
LOCUS CF138380
DEFINITION CF138380 UI-HF-BN0-aos-d-02-0-UI.r1 NIH MGC_50 Homo sapiens cDNA clone
IMAGE:3094658 5', mRNA sequence.
ACCESSION CF138380
VERSION CF138380.1 GI:33253824
KEYWORDS EST.

```

```

SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE        Normalization and subtraction: two approaches to facilitate gene
              discovery
JOURNAL       Genome Res. 6 (9), 791-806 (1996)
PUBMED
COMMENT      Contact: Soares, MB
              Coordinated Laboratory for Computational Genomics
              University of Iowa
              375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
              Tel: 319 335 8250
              Fax: 319 335 9585
              Email: bento-soares@uiowa.edu
              Tissue Procurement: Louis Staudt
              cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Distribution information can be found at
              http://genome.uiowa.edu/distribution/humanfl.html
              Seq primer: pYX-5.

FEATURES
source
Location/Qualifiers
1..426
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3094659"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LTI)"
/clone_lib="NIH MGC 50"
/note="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D. "

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 426;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20
|||||
Db 328 GCTTTGGTTGGGCAACACAT 309

RESULT 36
AA893671/c
LOCUS
DEFINITION EST197474 Normalized rat placenta, Bento Soares Rattus sp. cDNA
clone RPLA127 3' end, mRNA sequence.
ACCESSION AA893671
VERSION AA893671.1 GI:4132319
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Rattus.
1 (bases 1 to 427)
Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat
Gene Index
Unpublished (1998)
On Apr 3, 1998 this sequence version replaced gi:3020550.
Contact: Lee, NH

The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.

FEATURES
source
Location/Qualifiers
1..427
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="taxon:10118"
/clone="RPLA127"
/clone_lib="Normalized rat placenta, Bento Soares"
/note="Organ: placenta; Vector: pT7T3pac; Site_1: EcoRI;
Site_2: NotI"

ORIGIN
Query Match 100.0%; Score 20; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20
|||||
Db 262 GCTTTGGTTGGGCAACACAT 243

RESULT 37
AA124874
LOCUS
DEFINITION mp73508.r1 Soares thymus_2NDMT Mus musculus cDNA clone IMAGE:574862
5', similar to gb:U02368 PAIRED BOX PROTEIN PAX-3 (HUMAN);, mRNA
sequence.
ACCESSION AA124874
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
1 (bases 1 to 427)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through ILNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:349510
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 231.

FEATURES
source
Location/Qualifiers
1..427
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:574862"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares thymus_2NDMT"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

```

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pVT3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fátima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 427;  
Best Local Similarity 100.0%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAACACAT 20

Db 159 GCTTTGGTTGGGCAACACAT 178

## RESULT 38

BF414932

LOCUS

DEFINITION BF414932 433 bp mRNA linear EST 28-NOV-2000

UI-R-BJ2-bot-g-01-0-UI.s1 UI-R-BJ2 Rattus norvegicus cDNA clone

UI-R-BJ2-bot-g-01-0-UI 3', mRNA sequence.

ACCESSION BF414932

VERSION BF414932.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized AV canal at 16.5 dpc library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com). The following repetitive elements were found in this cDNA sequence: 1-21,

>AT\_rich#Low\_complexity

Seq primer: M13 Forward

POLYA=Yes.

## FEATURES

source

1. 433

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-BJ2-bot-g-01-0-UI"

/lab\_host="DHI10B (Life Technologies)"

/clone\_lib="UI-R-BJ2"

/note="Vector: pVT3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ2

library is a subcloned library derived from the following

tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc,

atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15

dpc, AV canal at 15 dpc. For a detailed description of

the library from which this clone was derived, please visit our web site at [rctest.eng.uiowa.edu](http://rctest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996) TAG TISSUE=AV canal at 16.5 dpc TAG LIB=UI-R-BJ2 TAG\_SEQ=GAACC"

## ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 433;  
Best Local Similarity 100.0%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGGCAACACAT 20

Db 173 GCTTTGGTTGGGCAACACAT 192

## RESULT 39

CA567706/c

LOCUS

DEFINITION

CA567706

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other ESTs: K0417G10-3

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: [cdna@gsun.grc.nia.nih.gov](mailto:cdna@gsun.grc.nia.nih.gov)

Plate: cDNA row: G column: 10

Seq primer: M13 Reverse

High quality sequence stop: 434

POLYA=No.

Location/Qualifiers

1. 434

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C3H/He"

/db\_xref="niaEST:K0417G10-5N"

/db\_xref="taxon:10090"

/clone="NIA:K0417G10 IMAGE:30060753"

/tissue\_type="Mesenchymal stem cell"

/cell\_line="9-15-C cells"

/lab\_host="DHI10B"

/clone\_lib="NIA Mouse Mesenchymal Stem Cell cDNA Library

(Long)"

/note="Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:

NotI; Mouse cDNA project by the Laboratory of Genetics,

National Institute on Aging (NIA), Intramural Research

Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is

a long-transcript enriched cDNA library (Ref. Genome Res.

11: 1553-1558 (2001). [PMID: 11544199]. Total RNAs were

obtained from Dr. Akihiro Umezawa (Keio University School

of Medicine, Japan). Double-stranded cDNAs were

synthesized with an Oligo(dT) primer [Invitrogen:

5'-pGATGATGTCGATCGAGCGCGCCCTTTT-3'] from

2.2 ug of total RNA, treated with T4 DNA polymerase, and

purified by ethanol-precipitation. The cDNAs were ligated

to Lone-linker L1-Sal4, purified by phenol/chloroform, and

separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 434;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||

Db 65 GCTTTGGTTGGCAACACAT 46  
|||||

RESULT 40  
AW912853 442 bp mRNA linear EST 25-MAY-2000  
LOCUS uf45g12.y1 Soares\_mammary\_gland\_NMLMG Mus musculus cDNA clone  
DEFINITION IMAGE:1514374 5' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3  
(HUMAN); mRNA sequence.

ACCESSION AW912853  
VERSION AW912853.1 GI:8078479  
KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
TITLE Sciurognathi; Muroidae; Muridae; Murinae; Mus.

JOURNAL 1 (bases 1 to 442)  
COMMENT NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:941226  
Seq primer: -40RP from Gibco  
High quality sequence stop: 406.

## FEATURES

source  
Location/Qualifiers  
1..442  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1514374"  
/sex="female (lactating)"  
/tissue\_type="mammary gland"  
/lab\_host="DH10B"  
/clone\_lib="Soares mammary gland\_NMLMG"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from mammary  
gland tissue from a lactating female, and was then primed  
with a Not I - oligo(dT) primer. Double-stranded cDNA was  
ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the  
modified pT73 vector. Library is normalized. Library  
was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 442;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||

Db 132 GCTTTGGTTGGCAACACAT 151

## RESULT 41

## AW280923

## LOCUS

## DEFINITION

z997b12.r1 NCI\_CGAP GCB1 Homo sapiens cDNA clone IMAGE:711551 5',  
similar to gb:U02368 PAIRED BOX PROTEIN PAX-3 (HUMAN); mRNA  
sequence.

## ACCESSION

## AW280923

## VERSION

## AW280923.1 GI:1923621

## KEYWORDS

## SOURCE

## ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 447)  
AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 408.

## FEATURES

source  
Location/Qualifiers  
1..447  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:711551"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP GCB1"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD-),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTACCAATCTGACGTGGAGCGCGCTCACTTTTCTTTT-3'  
]. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 447;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## Qy

1 GCTTTGGTTGGCAACACAT 20  
|||||

## Db

159 GCTTTGGTTGGCAACACAT 178  
|||||

## RESULT 42

## AA897422

## LOCUS

## DEFINITION

al48c09.s1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone  
IMAGE:1460560 3' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3  
(HUMAN); contains LTR1.b3 LTR1 repetitive element ; mRNA sequence.

## ACCESSION

## AA897422

## VERSION

## AA897422.1 GI:3034042

## KEYWORDS

## SOURCE

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 447)  
AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:941226  
Seq primer: -40RP from Gibco  
High quality sequence stop: 406.

## FEATURES

source  
Location/Qualifiers  
1..442  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1514374"  
/sex="female (lactating)"  
/tissue\_type="mammary gland"  
/lab\_host="DH10B"  
/clone\_lib="Soares mammary gland\_NMLMG"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from mammary  
gland tissue from a lactating female, and was then primed  
with a Not I - oligo(dT) primer. Double-stranded cDNA was  
ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the  
modified pT73 vector. Library is normalized. Library  
was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 442;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## Qy

1 GCTTTGGTTGGCAACACAT 20  
|||||

## Db

159 GCTTTGGTTGGCAACACAT 178  
|||||

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
1 (bases 1 to 448)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1331 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 443.
Location/Qualifiers
1. .448
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1460560"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CCAP GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match 100.0%; Score 20; DB 1; Length 448;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20
|||||
Db 160 GCTTTGGTTGGGCAACACAT 179

RESULT 43
BE849701 452 bp mRNA linear EST 26-SEP-2000
LOCUS uw88all.y1 Soares mammary_gland NMLMG Mus musculus cDNA clone
DEFINITION IMAGE:3469052 5' similar to gb:U02368 PAIRED BOX PROTEIN PAX-3
(HUMAN); mRNA sequence.
ACCESSION BE849701.1 GI:10308040
VERSION BE849701
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 452)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1388412
Seq primer: -40RP from Gibco
High quality sequence stop: 395.
Location/Qualifiers
1. .452
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3469052"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares mammary_gland NMLMG"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match 100.0%; Score 20; DB 2; Length 452;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20
|||||
Db 185 GCTTTGGTTGGGCAACACAT 204

RESULT 44
CO715537 453 bp mRNA linear EST 27-JUL-2004
LOCUS DG14-254d4 DG14-muscle Canis familiaris cDNA 3', mRNA sequence.
DEFINITION CO715537
ACCESSION CO715537
VERSION CO715537.1 GI:50687616
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
1 (bases 1 to 453)
Schlueter, T., Hermanns, J., Weindel, M., Schuette, D., Kranz, H.,
Henrich, J. and Loebbert, R.
DOG array cDNA clone collection
Unpublished (2004)
Contact: Thomas Schlueter
LION bioscience AG
Waldhoferstrasse 98, D-69123 Heidelberg, Germany
Tel: +49 6221 4038 150
Fax: +49 6221 4038 290
Email: Thomas.Schlueter@lionbioscience.com.
Location/Qualifiers
1. .453
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="Beagle"
/db_xref="taxon:9615"
/tissue_type="muscle"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="DG14-muscle"
/note="Organ: muscle; Vector: Dog pBluescript LION"

ORIGIN
Query Match 100.0%; Score 20; DB 7; Length 453;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAACACAT 20
|||||
Db 179 GCTTTGGTTGGGCAACACAT 198

RESULT 45

```

BI294032  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.  
 1 (bases 1 to 457)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 8899548  
 Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to identify it as a clone from the  
 normalized rat placenta pool library cDNA Library Preparation: M.B.  
 Soares Lab Clone Distribution: clones will be available through  
 Research Genetics (www.resgen.com) The following repetitive  
 elements were found in this cDNA sequence: 1-21,  
 >AT\_richLow complexity  
 Seq primer: M13 Forward  
 POLYA=Yes.

#### FEATURES

Location/Qualifiers  
 1..457  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clones="UI-R-DKO-cdy-e-08-0-UI"  
 /dev\_stage="ADULT"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="UI-R-DKO"  
 /note="Vector: p773D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-DKO  
 library is a subtracted library derived from a mixture of  
 five individually tagged normalized rat libraries:  
 brain-nRBP (20%), heart-nRHP (20%), kidney-nRKP (20%),  
 aorta-nRAP (20%), and placenta-nRPP (20%). Each original  
 library was constructed from a mixture of equal amounts of  
 RNA from seven different developmental time-points:  
 embryonic day 17, embryonic day 19, embryonic day 21,  
 adult day 1, adult day 12, adult day 75, and adult day  
 200. (Exception: the aorta pool does not contain embryonic  
 day 17 RNA and the placenta pool contains only the three  
 embryonic stages). Each library was normalized  
 individually according to the procedure described by  
 Bonaldo, Lennon & Soares (Genome Research Genome 6:  
 791-806, 1996). For construction of the DKO subtracted  
 library, plasmid DNA from each of the five individually  
 tagged normalized libraries was mixed in the proportions  
 specified above and electroporated into competent bacteria  
 for production of single-stranded circular DNA  
 representing the pool of libraries. Single-stranded  
 circular DNA representing these five normalized libraries  
 was then used as a tracer in a subtractive hybridization  
 with a driver (PCR amplified inserts from a plasmid DNA  
 template preparation) comprising: a) a set of about 1,000  
 arrayed clones from each of the five non-normalized

libraries of brain (CT0s), heart (CS0s), kidney (CU0s),  
 aorta (CW0s), and placenta (CX0s). The resulting pool of  
 approximately 5,000 clones represented about 33.3% of the  
 final driver population. A set of about 2,000 arrayed  
 clones from each of the five normalized libraries of brain  
 (CT0), heart (CS0), kidney (CU0), aorta (CW0), and  
 placenta (CX0). The resulting pool of about 10,000 clones  
 represented about 66.6% of the final driver population.  
 TAG TISSUE=rat placenta pool  
 TAG\_LIB=UI-R-DKO  
 TAG\_SEQ=TCACGACAGT"

#### ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 457;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCTTTGGTTGGGCAACAT 20  
 ||||||||||||||||  
 Db 173 GCTTTGGTTGGGCAACAT 192

#### RESULT 46

BG068406  
 LOCUS  
 DEFINITION  
 H3065B05-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone  
 H3065B05 3', mRNA sequence.  
 ACCESSION  
 BG068406  
 VERSION  
 BG068406.2 GI:40016024  
 KEYWORDS  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridea; Muridae; Murinae; Mus.  
 1 (bases 1 to 461)  
 Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,  
 Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,  
 Wood,W.H. III, Becker,K.G. and Ko,M.S.H.  
 Genome-wide expression profiling of mid-gestation placenta and  
 embryo using a 15,000 mouse developmental cDNA microarray  
 Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)  
 10922068  
 On Jan 26, 2001 this sequence version replaced gi:12550975.  
 Other ESTs: H3065B05-5  
 Contact: George J. Kargul  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@lgsun.grc.nia.nih.gov  
 This clone set has been freely distributed to the community. Please  
 visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.  
 Plate: H3065 ROW: B column: 05  
 Seq primer: -21M13 Forward  
 High quality sequence stop: 461  
 POLYA=Yes.

#### FEATURES

Location/Qualifiers  
 1..461  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="niaEST:H3065B05-3"  
 /db\_xref="taxon:10090"  
 /clone="H3065B05"  
 /sex="Clones arrayed from a variety of cDNA libraries"  
 /dev\_stage="Clones arrayed from a variety of cDNA libraries"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse 15K cDNA Clone Set"  
 /notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This  
 clone is among a rearrayed set of 15,247 clones from 11  
 embryo cDNA libraries (including preimplantation stage  
 embryos from unfertilized egg to blastocyst, embryonic

part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

## ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20  
 |||||  
 DB 156 GCTTTGGTTGGCAACACAT 175

## RESULT 47

AU022008  
 LOCUS 467 bp mRNA linear EST 20-OCT-1998  
 DEFINITION AU022008 Mouse unfertilized egg cDNA Mus musculus cDNA clone J0406F11 3', mRNA sequence.  
 ACCESSION AU022008  
 VERSION AU022008.1 GI:3386967  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
 1 (bases 1 to 467)  
 Ko,M.S.H, Kitchen,J.R., Wang,X., Threat,T.A., Sun,T., Depalma,G.E., Liang,Y., Kargul,G.J., Shaxara,R. and Doi,H.  
 Systematic analyses of genes expressed in unfertilized mouse eggs (The ERATO/Doi Project at Wayne State University) (Ko,M.S.H. et al.)  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Hirofumi Doi  
 Doi Biosymmetry Project, ERATO  
 Japan Science and Technology Corporation (JST)  
 WBG Marive East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan  
 Email: hdo@bio.jst.go.jp.  
 Location/Qualifiers  
 1..467  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CS7BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="J0406F11"  
 /dev\_stage="unfertilized egg"  
 /clone\_lib="Mouse unfertilized egg cDNA"

## FEATURES

## source

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20  
 |||||  
 DB 156 GCTTTGGTTGGCAACACAT 175

## RESULT 48

AI373959

LOCUS AI373959 468 bp mRNA linear EST 16-FEB-1999  
 DEFINITION Q291d11.x1 Soares pregnant\_uterus\_NbHPU Homo sapiens cDNA clone IMAGE:2041941 3', similar to gb:U02368 PAIRED BOX PROTEIN PAX-3 (HUMAN);, mRNA sequence.  
 ACCESSION AI373959  
 VERSION AI373959.1 GI:4153825  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 468)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs-r@mail.nih.gov  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 406 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 436.  
 Location/Qualifiers  
 1..468  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2041941"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares pregnant uterus NbHPU"  
 /note="Organ: uterus; Vector: pT7T3-Pac; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', AACTGAGAAATTCGCGCCGCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 468;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACACAT 20  
 |||||  
 DB 159 GCTTTGGTTGGCAACACAT 178

## RESULT 49

BB702712/c  
 LOCUS BB702712 471 bp mRNA linear EST 11-OCT-2001  
 DEFINITION BB702712 RIKEN full-length enriched, in vitro fertilized eggs Mus musculus cDNA clone 7420443L11 3', mRNA sequence.  
 ACCESSION BB702712  
 VERSION BB702712.1 GI:16051536  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
 1 (bases 1 to 471)  
 Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,

Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tomaru,A., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.  
2001)

## TITLE

## JOURNAL

## COMMENT

Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suicho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,  
Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.  
and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,  
Sugahara,Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

## e mouse tissues.

## FEATURES

## source

Location/Qualifiers  
1. .471  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="7420443L11"  
/sex="female"  
/tissue\_type="in vitro fertilized eggs"  
/dev\_stage="egg"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, in vitro  
fertilized eggs"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. Second strand cDNA was prepared with the  
primer adapter of sequence [5',  
GAGAGAGAGATTCGAGTTAATTAATTAATCCGCCGCC 3']. cDNA  
was cloned into the XhoI and BamHI sites. Vector: a  
modified pBluescript KS(+) after bulk excision from Lambda  
FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

## ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 471;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20

|||||  
Db 316 GCTTTGGTTGGCAACACAT 297

## RESULT 50

## BF388801

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## PUBMED

## COMMENT

BF388801 472 bp mRNA linear EST 27-NOV-2000  
UI-R-BS2-bde-f-12-0-UI.s1 UI-R-BS2 Rattus norvegicus cDNA clone  
UI-R-BS2-bde-f-12-0-UI 3', mRNA sequence.

BF388801

BF388801.1 GI:11373625

EST.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Rattus norvegicus

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognath; Muroidae; Muridae; Murinae; Rattus.

1 (bases 1 to 472)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the

normalized embryo at 13 dpc library cDNA Library Preparation: M.B.

Soares Lab Clone distribution: clones will be available through

Research Genetics ([www.resgen.com](http://www.resgen.com))

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1. .472

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-BS2-bde-f-12-0-UI"

/dev\_stage="embryonic 13 dpc"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="UI-R-BS2"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BS2

library is a subcloned library derived from 13 dpc whole

embryo tissue. For a detailed description of the library

from which this clone was derived, please visit our web

site at [ratest.eng.uiowa.edu](http://ratest.eng.uiowa.edu). The subtraction has been

previously described in (Bonaldo, Lennon and Soares,

Genome Research 6:791-806, 1996)

TAG TISSUE=embryo at 13 dpc

TAG LIB=UI-R-BS2

TAG\_SEQ=AATCC"

Query Match 100.0%; Score 20; DB 2; Length 472;

Best Local Similarity 100.0%; Pred. No. 0.31;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20

|||||

Db 172 GCTTTGGTTGGCAACACAT 191

|||||

Search completed: May 8, 2006, 02:26:04

Job time : 3834 secs



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2006, 00:50:50 ; Search time 427 Seconds  
(without alignments)  
190.751 Million cell updates/sec

Title: US-10-671-074-172

Perfect score: 20

Sequence: 1 gcttggtggcaacacat 20

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 9306428 seqs, 2036268586 residues

Word size : 1

Total number of hits satisfying chosen parameters: 18612072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA New:

- 1: /SIDSS/ptodata/1/pubpna/US08 NEW PUB.seq1.\*
- 2: /SIDSS/ptodata/1/pubpna/US06 NEW PUB.seq.\*
- 3: /SIDSS/ptodata/1/pubpna/US07 NEW PUB.seq.\*
- 4: /SIDSS/ptodata/1/pubpna/US08 NEW PUB.seq.\*
- 5: /SIDSS/ptodata/1/pubpna/PCT NEW PUB.seq.\*
- 6: /SIDSS/ptodata/1/pubpna/US09 NEW PUB.seq.\*
- 7: /SIDSS/ptodata/1/pubpna/US09 NEW PUB.seq1.\*
- 8: /SIDSS/ptodata/1/pubpna/US09 NEW PUB.seq2.\*
- 9: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq.\*
- 10: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq1.\*
- 11: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq2.\*
- 12: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq3.\*
- 13: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq4.\*
- 14: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq.\*
- 15: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq1.\*
- 16: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq2.\*
- 17: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq3.\*
- 18: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq4.\*
- 19: /SIDSS/ptodata/1/pubpna/US60 NEW PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Ref. No.	Score	Match	Length	ID	Description
C 1	20	100.0	25	17	US-11-136-527-226522
C 2	20	100.0	254	12	US-10-920-152-12
C 3	20	100.0	600	17	US-11-136-527-4790
C 4	20	100.0	1803	17	US-11-136-527-3494
C 5	20	100.0	3517	13	US-10-469-469-192
C 6	19	95.0	25	17	US-11-136-527-34022
C 7	19	95.0	25	17	US-11-136-527-226492
C 8	19	95.0	25	17	US-11-136-527-226494
C 9	18	90.0	25	17	US-11-136-527-34046
C 10	17	85.0	25	17	US-11-136-527-34023
C 11	16	80.0	25	17	US-11-136-527-34042
C 12	16	80.0	995	12	US-10-301-480-546810
C 13	16	80.0	995	12	US-10-301-480-1160219
C 14	16	80.0	996	12	US-10-301-480-553142

C 15	80.0	16	12	US-10-301-480-553488	Sequence 553488,
C 16	80.0	16	12	US-10-301-480-1168551	Sequence 1168551,
C 17	80.0	16	12	US-10-301-480-1168897	Sequence 1168897,
C 18	75.0	25	17	US-11-136-527-34024	Sequence 34024, A
C 19	75.0	1000	12	US-10-301-480-541667	Sequence 541667,
C 20	75.0	1000	12	US-10-301-480-1155076	Sequence 1155076,
C 21	75.0	1153	7	US-09-925-065A-669805	Sequence 669805,
C 22	75.0	3593	11	US-10-115-609-22	Sequence 22, Appl
C 23	70.0	25	17	US-11-136-527-34011	Sequence 34011, A
C 24	70.0	201	17	US-11-124-367A-25760	Sequence 25760, A
C 25	70.0	403	7	US-09-925-065A-153801	Sequence 153801,
C 26	70.0	404	12	US-10-301-480-247396	Sequence 247396,
C 27	70.0	404	12	US-10-301-480-860805	Sequence 860805,
C 28	70.0	468	11	US-10-301-480-14320	Sequence 14320, A
C 29	70.0	468	12	US-10-301-480-627729	Sequence 627729,
C 30	70.0	506	7	US-09-925-065A-603534	Sequence 603534,
C 31	70.0	515	7	US-09-925-065A-603533	Sequence 603533,
C 32	70.0	535	12	US-10-301-480-425083	Sequence 425083,
C 33	70.0	535	12	US-10-301-480-425084	Sequence 425084,
C 34	70.0	535	12	US-10-301-480-425085	Sequence 425085,
C 35	70.0	535	12	US-10-301-480-1038492	Sequence 1038492,
C 36	70.0	535	12	US-10-301-480-1038493	Sequence 1038493,
C 37	70.0	535	12	US-10-301-480-1038494	Sequence 1038494,
C 38	70.0	537	7	US-09-925-065A-353424	Sequence 353424,
C 39	70.0	537	7	US-09-925-065A-353425	Sequence 353425,
C 40	70.0	537	7	US-09-925-065A-353426	Sequence 353426,
C 41	70.0	538	7	US-09-925-065A-603532	Sequence 603532,
C 42	70.0	545	7	US-09-925-065A-94527	Sequence 94527, A
C 43	70.0	545	11	US-10-301-480-195769	Sequence 195769,
C 44	70.0	545	12	US-10-301-480-809178	Sequence 809178,
C 45	70.0	556	12	US-10-301-480-415497	Sequence 415497,
C 46	70.0	556	12	US-10-301-480-415498	Sequence 415498,
C 47	70.0	556	12	US-10-301-480-1028906	Sequence 1028906,
C 48	70.0	556	12	US-10-301-480-1028907	Sequence 1028907,
C 49	70.0	562	7	US-09-925-065A-386479	Sequence 386479,
C 50	70.0	564	12	US-10-301-480-455055	Sequence 455055,
C 51	70.0	564	12	US-10-301-480-1068464	Sequence 1068464,
C 52	70.0	567	7	US-09-925-065A-342764	Sequence 342764,
C 53	70.0	567	7	US-09-925-065A-342765	Sequence 342765,
C 54	70.0	587	7	US-09-925-065A-605689	Sequence 605689,
C 55	70.0	592	7	US-09-925-065A-199507	Sequence 199507,
C 56	70.0	592	7	US-09-925-065A-463575	Sequence 463575,
C 57	70.0	592	7	US-09-925-065A-463576	Sequence 463576,
C 58	70.0	592	7	US-09-925-065A-463577	Sequence 463577,
C 59	70.0	592	7	US-09-925-065A-463578	Sequence 463578,
C 60	70.0	593	12	US-10-301-480-288200	Sequence 288200,
C 61	70.0	593	12	US-10-301-480-901609	Sequence 901609,
C 62	70.0	593	12	US-10-301-480-520049	Sequence 520049,
C 63	70.0	595	12	US-10-301-480-520050	Sequence 520050,
C 64	70.0	595	12	US-10-301-480-520051	Sequence 520051,
C 65	70.0	595	12	US-10-301-480-520052	Sequence 520052,
C 66	70.0	595	12	US-10-301-480-1133458	Sequence 1133458,
C 67	70.0	595	12	US-10-301-480-1133459	Sequence 1133459,
C 68	70.0	595	12	US-10-301-480-1133460	Sequence 1133460,
C 69	70.0	600	10	US-10-301-480-1133461	Sequence 1133461,
C 70	70.0	600	10	US-10-750-185-3665	Sequence 3665, Ap
C 71	70.0	601	7	US-09-925-065A-3665	Sequence 3665, Ap
C 72	70.0	602	7	US-09-925-065A-540729	Sequence 540729,
C 73	70.0	606	7	US-09-925-065A-902618	Sequence 902618,
C 74	70.0	611	7	US-09-925-065A-199508	Sequence 199508,
C 75	70.0	661	7	US-09-925-065A-884106	Sequence 884106,
C 76	70.0	687	18	US-11-079-463-5066	Sequence 884132,
C 77	70.0	733	12	US-10-301-480-578772	Sequence 5066, Ap
C 78	70.0	733	12	US-10-301-480-1192181	Sequence 578772,
C 79	70.0	998	12	US-10-301-480-536665	Sequence 1192181,
C 80	70.0	998	12	US-10-301-480-536665	Sequence 536665,
C 81	70.0	1209	16	US-11-082-389-417	Sequence 1150074,
C 82	70.0	1335	11	US-10-301-480-27409	Sequence 417, App
C 83	70.0	1335	11	US-10-301-480-27410	Sequence 27409, A
C 84	70.0	1335	11	US-10-301-480-27411	Sequence 27410, A
C 85	70.0	1335	11	US-10-301-480-27412	Sequence 27411, A
C 86	70.0	1335	11	US-10-301-480-27412	Sequence 27412, A
C 87	70.0	1335	12	US-10-301-480-640818	Sequence 640818,

c 88	14	70.0	1335	12	US-10-301-480-640819	Sequence 640819,	161	13	65.0	566	12	US-10-301-480-634114	Sequence 634114,
c 89	14	70.0	1335	12	US-10-301-480-640820	Sequence 640820,	162	13	65.0	588	7	US-09-925-065A-521903	Sequence 521903,
c 90	14	70.0	1335	12	US-10-301-480-640821	Sequence 640821,	163	13	65.0	588	12	US-10-301-480-408727	Sequence 408727,
c 91	14	70.0	1839	18	US-11-079-463-1731	Sequence 1731, Ap	164	13	65.0	588	12	US-10-301-480-1022136	Sequence 1022136,
c 92	14	70.0	1931	7	US-09-925-065A-723533	Sequence 723533,	165	13	65.0	590	7	US-09-925-065A-648051	Sequence 648051,
c 93	14	70.0	2196	7	US-09-925-065A-683172	Sequence 683172,	166	13	65.0	591	7	US-09-925-065A-648059	Sequence 648059,
c 94	14	70.0	2197	13	US-10-469-469-213	Sequence 213, App	c 167	13	65.0	591	7	US-09-925-065A-935924	Sequence 935924,
c 95	14	70.0	2197	13	US-10-469-469-215	Sequence 215, App	c 168	13	65.0	591	7	US-09-925-065A-935925	Sequence 935925,
c 96	14	70.0	2314	11	US-10-301-480-96382	Sequence 96382, A	c 169	13	65.0	592	7	US-09-925-065A-33272	Sequence 33272,
c 97	14	70.0	2314	11	US-10-301-480-96383	Sequence 96383, A	c 170	13	65.0	596	7	US-09-925-065A-754854	Sequence 754854,
c 98	14	70.0	2314	12	US-10-301-480-709791	Sequence 709791,	c 171	13	65.0	596	7	US-09-925-065A-754855	Sequence 754855,
c 99	14	70.0	2314	12	US-10-301-480-709792	Sequence 709792,	c 172	13	65.0	597	7	US-09-925-065A-825513	Sequence 825513,
c 100	14	70.0	2493	11	US-10-932-182A-78335	Sequence 78335, A	c 173	13	65.0	597	11	US-10-301-480-9010	Sequence 9010, Ap
c 101	14	70.0	2493	11	US-10-932-182A-78335	Sequence 78335, A	c 174	13	65.0	597	11	US-10-301-480-622419	Sequence 622419, Ap
c 102	14	70.0	3160	18	US-11-141-611-51	Sequence 51, Appl	c 175	13	65.0	600	10	US-10-750-185-2706	Sequence 2706, Ap
c 103	14	70.0	3256	7	US-09-925-065A-688254	Sequence 688254,	c 176	13	65.0	600	10	US-10-750-623-2706	Sequence 2706, Ap
c 104	14	70.0	52192	10	US-10-395-561-13231	Sequence 13231, A	c 177	13	65.0	609	7	US-09-925-065A-478182	Sequence 478182,
c 105	13	65.0	18	10	US-10-310-914A-171448	Sequence 171448,	c 178	13	65.0	609	7	US-09-925-065A-478183	Sequence 478183,
c 106	13	65.0	19	10	US-10-310-914A-1371526	Sequence 1371526,	c 179	13	65.0	609	11	US-10-301-480-58560	Sequence 58560, A
c 107	13	65.0	21	10	US-10-310-914A-171416	Sequence 171416,	c 180	13	65.0	609	11	US-10-301-480-58561	Sequence 58561, A
c 108	13	65.0	21	10	US-10-310-914A-791926	Sequence 791926,	c 181	13	65.0	609	12	US-10-301-480-671969	Sequence 671969,
c 109	13	65.0	22	10	US-10-310-914A-171417	Sequence 171417,	c 182	13	65.0	609	12	US-10-301-480-671970	Sequence 671970,
c 110	13	65.0	23	10	US-10-310-914A-171437	Sequence 171437,	c 183	13	65.0	616	7	US-09-925-065A-471520	Sequence 471520,
c 111	13	65.0	23	10	US-10-310-914A-791806	Sequence 791806,	c 184	13	65.0	621	7	US-09-925-065A-481726	Sequence 481726,
c 112	13	65.0	23	10	US-10-310-914A-791830	Sequence 791830,	c 185	13	65.0	621	7	US-09-925-065A-481727	Sequence 481727,
c 113	13	65.0	24	10	US-10-310-914A-171425	Sequence 171425,	c 186	13	65.0	624	11	US-10-932-182A-5628	Sequence 5628, Ap
c 114	13	65.0	24	10	US-10-310-914A-1371594	Sequence 1371594,	c 187	13	65.0	624	11	US-10-932-182A-5628	Sequence 5628, Ap
c 115	13	65.0	25	11	US-10-932-182A-116229	Sequence 116229,	c 188	13	65.0	631	7	US-09-925-065A-411439	Sequence 411439,
c 116	13	65.0	25	11	US-10-932-182A-116229	Sequence 116229,	c 189	13	65.0	638	7	US-09-925-065A-49102	Sequence 49102, A
c 117	13	65.0	25	11	US-10-934-048A-3168	Sequence 3168, Ap	c 190	13	65.0	638	11	US-10-301-480-150340	Sequence 150340, A
c 118	13	65.0	25	17	US-11-121-849-216127	Sequence 216127,	c 191	13	65.0	638	12	US-10-301-480-763749	Sequence 763749,
c 119	13	65.0	25	17	US-11-136-527-34009	Sequence 34009, A	c 192	13	65.0	640	12	US-10-301-480-477647	Sequence 477647,
c 120	13	65.0	25	17	US-11-136-527-226524	Sequence 226524,	c 193	13	65.0	640	12	US-10-301-480-1091056	Sequence 1091056,
c 121	13	65.0	201	10	US-10-995-561-50021	Sequence 50021, A	c 194	13	65.0	647	7	US-09-925-065A-512632	Sequence 512632,
c 122	13	65.0	377	7	US-09-925-065A-491370	Sequence 491370,	c 195	13	65.0	647	7	US-09-925-065A-512633	Sequence 512633,
c 123	13	65.0	425	7	US-09-925-065A-459665	Sequence 459665,	c 196	13	65.0	655	7	US-09-925-065A-763741	Sequence 763741,
c 124	13	65.0	429	12	US-10-301-480-517007	Sequence 517007,	c 197	13	65.0	655	7	US-09-925-065A-763742	Sequence 763742,
c 125	13	65.0	429	12	US-10-301-480-517007	Sequence 517007,	c 198	13	65.0	656	7	US-09-925-065A-771621	Sequence 771621,
c 126	13	65.0	454	18	US-11-116-881A-1130416	Sequence 1130416,	c 199	13	65.0	656	7	US-09-925-065A-771622	Sequence 771622,
c 127	13	65.0	506	17	US-11-128-061-6409	Sequence 6409, Ap	c 200	13	65.0	658	7	US-09-925-065A-882868	Sequence 882868,
c 128	13	65.0	506	17	US-11-128-061-6409	Sequence 6409, Ap	c 201	13	65.0	658	7	US-09-925-065A-889432	Sequence 889432,
c 129	13	65.0	511	7	US-09-925-065A-60422	Sequence 60422, A	c 202	13	65.0	670	12	US-10-301-480-562512	Sequence 562512,
c 130	13	65.0	511	12	US-10-301-480-161660	Sequence 161660,	c 203	13	65.0	670	12	US-10-301-480-1175921	Sequence 1175921,
c 131	13	65.0	511	12	US-10-301-480-775069	Sequence 775069,	c 204	13	65.0	683	7	US-09-925-065A-79305	Sequence 79305, A
c 132	13	65.0	515	7	US-09-925-065A-590910	Sequence 590910,	c 205	13	65.0	683	11	US-10-301-480-180544	Sequence 180544,
c 133	13	65.0	520	7	US-09-925-065A-545704	Sequence 545704,	c 206	13	65.0	683	12	US-10-301-480-793953	Sequence 793953,
c 134	13	65.0	520	7	US-09-925-065A-545705	Sequence 545705,	c 207	13	65.0	700	12	US-10-301-480-50992	Sequence 50992,
c 135	13	65.0	521	12	US-10-301-480-309528	Sequence 309528,	c 208	13	65.0	700	12	US-10-301-480-1164401	Sequence 1164401,
c 136	13	65.0	521	12	US-10-301-480-309528	Sequence 309528,	c 209	13	65.0	713	7	US-09-925-065A-82461	Sequence 82461, A
c 137	13	65.0	528	7	US-09-925-065A-224786	Sequence 224786,	c 210	13	65.0	713	11	US-10-301-480-183701	Sequence 183701,
c 138	13	65.0	535	17	US-11-128-061-2767	Sequence 2767, Ap	c 211	13	65.0	713	12	US-10-301-480-797110	Sequence 797110,
c 139	13	65.0	535	17	US-11-128-049-2767	Sequence 2767, Ap	c 212	13	65.0	725	7	US-09-925-065A-80813	Sequence 80813, A
c 140	13	65.0	542	7	US-09-925-065A-755300	Sequence 755300,	c 213	13	65.0	725	7	US-09-925-065A-80814	Sequence 80814, A
c 141	13	65.0	545	7	US-09-925-065A-471519	Sequence 471519,	c 214	13	65.0	725	11	US-10-301-480-182052	Sequence 182052,
c 142	13	65.0	545	7	US-09-925-065A-471521	Sequence 471521,	c 215	13	65.0	725	11	US-10-301-480-182053	Sequence 182053,
c 143	13	65.0	545	7	US-09-925-065A-471522	Sequence 471522,	c 216	13	65.0	725	11	US-10-301-480-182053	Sequence 182053,
c 144	13	65.0	546	7	US-09-925-065A-95641	Sequence 95641, A	c 217	13	65.0	725	12	US-10-301-480-795461	Sequence 795461,
c 145	13	65.0	546	7	US-09-925-065A-95642	Sequence 95642, A	c 218	13	65.0	736	7	US-09-925-065A-932077	Sequence 932077,
c 146	13	65.0	547	7	US-09-925-065A-507101	Sequence 507101,	c 219	13	65.0	736	7	US-09-925-065A-932077	Sequence 932077,
c 147	13	65.0	554	12	US-10-301-480-407288	Sequence 407288,	c 220	13	65.0	736	7	US-09-925-065A-842276	Sequence 842276,
c 148	13	65.0	554	12	US-10-301-480-459970	Sequence 459970,	c 221	13	65.0	736	11	US-10-301-480-185466	Sequence 185466,
c 149	13	65.0	554	12	US-10-301-480-1020697	Sequence 1020697,	c 222	13	65.0	736	11	US-10-301-480-798875	Sequence 798875,
c 150	13	65.0	554	12	US-10-301-480-1020697	Sequence 1020697,	c 223	13	65.0	736	11	US-10-301-480-586107	Sequence 586107,
c 151	13	65.0	555	7	US-09-925-065A-333718	Sequence 333718,	c 224	13	65.0	736	12	US-10-301-480-586107	Sequence 586107,
c 152	13	65.0	558	7	US-09-925-065A-640608	Sequence 640608,	c 225	13	65.0	788	7	US-09-925-065A-69475	Sequence 69475, A
c 153	13	65.0	559	7	US-09-925-065A-391767	Sequence 391767,	c 226	13	65.0	788	11	US-10-301-480-170714	Sequence 170714,
c 154	13	65.0	561	12	US-10-301-480-59681	Sequence 59681, A	c 227	13	65.0	788	11	US-10-301-480-170715	Sequence 170715,
c 155	13	65.0	561	12	US-10-301-480-673090	Sequence 673090,	c 228	13	65.0	788	12	US-10-301-480-784123	Sequence 784123,
c 156	13	65.0	563	11	US-10-301-480-196812	Sequence 196812,	c 229	13	65.0	788	12	US-10-301-480-784123	Sequence 784123,
c 157	13	65.0	563	11	US-10-301-480-196813	Sequence 196813,	c 230	13	65.0	788	12	US-10-301-480-784124	Sequence 784124,
c 158	13	65.0	563	12	US-10-301-480-810221	Sequence 810221,	c 231	13	65.0	794	10	US-10-750-185-55483	Sequence 55483, A
c 159	13	65.0	563	12	US-10-301-480-810222	Sequence 810222,	c 232	13	65.0	810	7	US-09-925-065A-86355	Sequence 86355, A
c 160	13	65.0	566	11	US-10-301-480-20705	Sequence 20705, A	c 233	13	65.0	810	11	US-10-301-480-187595	Sequence 187595,

234	13	65.0	810	12	US-10-301-480-801004	Sequence 801004, A	307	13	65.0	2849	10	US-10-750-185-51387	Sequence 51387, A
235	13	65.0	815	10	US-10-750-185-34762	Sequence 34762, A	308	13	65.0	2849	10	US-10-750-623-51387	Sequence 51387, A
236	13	65.0	815	10	US-10-750-623-34762	Sequence 34762, A	309	13	65.0	2905	7	US-09-925-065A-712237	Sequence 712237, A
237	13	65.0	838	12	US-10-301-480-600595	Sequence 600595, A	310	13	65.0	2905	7	US-09-925-065A-712238	Sequence 712238, A
238	13	65.0	838	12	US-10-301-480-1214004	Sequence 1214004, A	311	13	65.0	2905	7	US-09-925-065A-712239	Sequence 712239, A
239	13	65.0	876	7	US-09-925-065A-715791	Sequence 715791, A	312	13	65.0	2905	7	US-09-925-065A-712240	Sequence 712240, A
240	13	65.0	876	7	US-09-925-065A-715792	Sequence 715792, A	313	13	65.0	3000	17	US-11-194-246-277	Sequence 277, Appl
241	13	65.0	876	7	US-09-925-065A-715793	Sequence 715793, A	314	13	65.0	3024	7	US-09-925-065A-711172	Sequence 711172, A
242	13	65.0	971	12	US-10-301-480-565150	Sequence 565150, A	315	13	65.0	3024	7	US-09-925-065A-711173	Sequence 711173, A
243	13	65.0	971	12	US-10-301-480-1178559	Sequence 1178559, A	316	13	65.0	3139	10	US-10-750-185-26454	Sequence 26454, A
244	13	65.0	981	12	US-10-301-480-532124	Sequence 532124, A	317	13	65.0	3139	10	US-10-750-623-26454	Sequence 26454, A
245	13	65.0	981	12	US-10-301-480-1145533	Sequence 1145533, A	318	13	65.0	3601	7	US-09-925-065A-81422	Sequence 81422, A
246	13	65.0	993	12	US-10-301-480-534628	Sequence 534628, A	319	13	65.0	3601	7	US-09-925-065A-81423	Sequence 81423, A
247	13	65.0	993	12	US-10-301-480-1148037	Sequence 1148037, A	320	13	65.0	3601	7	US-09-925-065A-81424	Sequence 81424, A
248	13	65.0	998	12	US-10-301-480-565149	Sequence 565149, A	321	13	65.0	3601	7	US-09-925-065A-81425	Sequence 81425, A
249	13	65.0	998	12	US-10-301-480-1178558	Sequence 1178558, A	322	13	65.0	3601	11	US-10-301-480-182662	Sequence 182662, A
250	13	65.0	999	12	US-10-301-480-546090	Sequence 546090, A	323	13	65.0	3601	11	US-10-301-480-182663	Sequence 182663, A
251	13	65.0	999	12	US-10-301-480-1159499	Sequence 1159499, A	324	13	65.0	3601	11	US-10-301-480-182664	Sequence 182664, A
252	13	65.0	1052	10	US-10-750-185-49803	Sequence 49803, A	325	13	65.0	3601	11	US-10-301-480-182665	Sequence 182665, A
253	13	65.0	1052	10	US-10-750-623-49803	Sequence 49803, A	326	13	65.0	3601	12	US-10-301-480-796071	Sequence 796071, A
254	13	65.0	1092	17	US-11-166-609-5	Sequence 5, Appl	327	13	65.0	3601	12	US-10-301-480-796072	Sequence 796072, A
255	13	65.0	1118	7	US-09-925-065A-71947	Sequence 71947, A	328	13	65.0	3601	12	US-10-301-480-796073	Sequence 796073, A
256	13	65.0	1118	7	US-09-925-065A-71948	Sequence 71948, A	329	13	65.0	3601	12	US-10-301-480-796074	Sequence 796074, A
257	13	65.0	1118	11	US-10-301-480-173186	Sequence 173186, A	330	13	65.0	3897	17	US-11-166-609-7	Sequence 7, Appl
258	13	65.0	1118	11	US-10-301-480-173187	Sequence 173187, A	331	13	65.0	4190	17	US-11-091-883-430	Sequence 430, Appl
259	13	65.0	1118	12	US-10-301-480-786595	Sequence 786595, A	332	13	65.0	35962	10	US-10-775-169-234	Sequence 234, Appl
260	13	65.0	1118	12	US-10-301-480-786596	Sequence 786596, A	333	13	65.0	36297	17	US-11-124-367A-5023	Sequence 5023, Appl
261	13	65.0	1200	7	US-09-925-065A-2692	Sequence 2692, Appl	334	13	65.0	54946	10	US-10-995-561-13479	Sequence 13479, Appl
262	13	65.0	1200	7	US-09-925-065A-2693	Sequence 2693, Appl	335	13	65.0	79329	11	US-10-330-773-483	Sequence 483, Appl
263	13	65.0	1200	11	US-10-301-480-103929	Sequence 103929, A	336	13	65.0	100000	17	US-11-124-367A-5027	Sequence 5027, Appl
264	13	65.0	1200	11	US-10-301-480-103930	Sequence 103930, A	337	13	65.0	160213	17	US-11-121-086-103	Sequence 103, Appl
265	13	65.0	1200	12	US-10-301-480-717338	Sequence 717338, A	338	13	65.0	185393	17	US-11-121-086-101	Sequence 101, Appl
266	13	65.0	1200	12	US-10-301-480-717339	Sequence 717339, A	339	13	65.0	197241	18	US-11-114-798-47	Sequence 47, Appl
267	13	65.0	1304	11	US-10-301-480-38178	Sequence 38178, A	340	13	65.0	201990	10	US-10-995-561-13303	Sequence 13303, A
268	13	65.0	1304	12	US-10-301-480-651587	Sequence 651587, A	341	13	65.0	1691140	17	US-11-091-018-1	Sequence 1, Appl
269	13	65.0	1369	10	US-10-750-185-63004	Sequence 63004, A	342	13	65.0	2944528	18	US-11-045-004-1	Sequence 1, Appl
270	13	65.0	1369	10	US-10-750-623-63004	Sequence 63004, A	343	12	60.0	18	10	US-10-310-914A-582773	Sequence 582773, A
271	13	65.0	1383	10	US-10-750-185-48654	Sequence 48654, A	344	12	60.0	18	10	US-10-310-914A-582875	Sequence 582875, A
272	13	65.0	1383	10	US-10-750-623-48654	Sequence 48654, A	345	12	60.0	18	10	US-10-310-914A-1278546	Sequence 1278546, A
273	13	65.0	1415	11	US-10-301-480-159918	Sequence 159918, A	346	12	60.0	19	10	US-10-310-914A-582876	Sequence 582876, A
274	13	65.0	1418	7	US-09-925-065A-58680	Sequence 58680, A	347	12	60.0	19	10	US-10-310-914A-790893	Sequence 790893, A
275	13	65.0	1418	12	US-10-301-480-773327	Sequence 773327, A	348	12	60.0	19	15	US-11-101-244-619761	Sequence 619761, A
276	13	65.0	1422	10	US-10-750-185-58051	Sequence 58051, A	349	12	60.0	19	15	US-11-101-244-619858	Sequence 619858, A
277	13	65.0	1422	10	US-10-750-623-58051	Sequence 58051, A	350	12	60.0	19	15	US-11-101-244-1014343	Sequence 1014343, A
278	13	65.0	1451	18	US-11-096-568A-2439	Sequence 2439, Appl	351	12	60.0	19	15	US-11-101-244-1536666	Sequence 1536666, A
279	13	65.0	1467	11	US-10-932-182A-75635	Sequence 75635, A	352	12	60.0	19	16	US-11-083-784-619761	Sequence 619761, A
280	13	65.0	1467	11	US-10-932-182A-75635	Sequence 75635, A	353	12	60.0	19	16	US-11-083-784-619858	Sequence 619858, A
281	13	65.0	1584	11	US-10-932-182A-76285	Sequence 76285, A	354	12	60.0	19	16	US-11-083-784-1014343	Sequence 1014343, A
282	13	65.0	1584	11	US-10-932-182A-76285	Sequence 76285, A	355	12	60.0	19	16	US-11-083-784-1536666	Sequence 1536666, A
283	13	65.0	1860	10	US-10-750-185-59479	Sequence 59479, A	356	12	60.0	20	10	US-10-310-914A-582742	Sequence 582742, A
284	13	65.0	1860	10	US-10-750-623-59479	Sequence 59479, A	357	12	60.0	20	10	US-10-310-914A-582742	Sequence 582742, A
285	13	65.0	1888	7	US-09-925-065A-707588	Sequence 707588, A	358	12	60.0	21	10	US-10-310-914A-152877	Sequence 152877, A
286	13	65.0	2009	7	US-09-925-065A-82092	Sequence 82092, A	359	12	60.0	21	10	US-10-310-914A-582750	Sequence 582750, A
287	13	65.0	2009	7	US-09-925-065A-82093	Sequence 82093, A	360	12	60.0	21	10	US-10-310-914A-1278520	Sequence 1278520, A
288	13	65.0	2009	7	US-09-925-065A-82094	Sequence 82094, A	361	12	60.0	22	10	US-10-310-914A-41289	Sequence 41289, A
289	13	65.0	2009	11	US-10-301-480-183332	Sequence 183332, A	362	12	60.0	22	10	US-10-310-914A-152854	Sequence 152854, A
290	13	65.0	2009	11	US-10-301-480-183333	Sequence 183333, A	363	12	60.0	22	10	US-10-310-914A-1119639	Sequence 1119639, A
291	13	65.0	2009	11	US-10-301-480-183334	Sequence 183334, A	364	12	60.0	22	10	US-10-310-914A-1278536	Sequence 1278536, A
292	13	65.0	2009	12	US-10-301-480-796741	Sequence 796741, A	365	12	60.0	23	10	US-10-310-914A-136892	Sequence 136892, A
293	13	65.0	2009	12	US-10-301-480-796742	Sequence 796742, A	366	12	60.0	23	10	US-10-310-914A-550881	Sequence 550881, A
294	13	65.0	2009	12	US-10-301-480-796743	Sequence 796743, A	367	12	60.0	23	10	US-10-310-914A-1119616	Sequence 1119616, A
295	13	65.0	2083	7	US-09-925-065A-347048	Sequence 347048, A	368	12	60.0	23	10	US-10-310-914A-1278547	Sequence 1278547, A
296	13	65.0	2083	11	US-10-301-480-93373	Sequence 93373, A	369	12	60.0	24	10	US-10-310-914A-25888	Sequence 25888, A
297	13	65.0	2083	12	US-10-301-480-524998	Sequence 524998, A	370	12	60.0	24	10	US-10-310-914A-41290	Sequence 41290, A
298	13	65.0	2083	12	US-10-301-480-706782	Sequence 706782, A	371	12	60.0	24	10	US-10-310-914A-136883	Sequence 136883, A
299	13	65.0	2083	12	US-10-301-480-1138407	Sequence 1138407, A	372	12	60.0	24	10	US-10-310-914A-405646	Sequence 405646, A
300	13	65.0	2644	10	US-10-750-185-39558	Sequence 39558, A	373	12	60.0	24	10	US-10-310-914A-550877	Sequence 550877, A
301	13	65.0	2644	10	US-10-750-623-39558	Sequence 39558, A	374	12	60.0	24	10	US-10-310-914A-582877	Sequence 582877, A
302	13	65.0	2661	11	US-10-301-480-98679	Sequence 98679, A	375	12	60.0	24	10	US-10-310-914A-1119614	Sequence 1119614, A
303	13	65.0	2661	11	US-10-301-480-98679	Sequence 98679, A	376	12	60.0	24	10	US-10-310-914A-1119617	Sequence 1119617, A
304	13	65.0	2661	12	US-10-301-480-712087	Sequence 712087, A	377	12	60.0	25	10	US-10-310-914A-405610	Sequence 405610, A
305	13	65.0	2661	12	US-10-301-480-712088	Sequence 712088, A	378	12	60.0	25	10	US-10-310-914A-582904	Sequence 582904, A
306	13	65.0	2787	10	US-10-467-657-5655	Sequence 5655, Appl	379	12	60.0	25	10	US-10-310-914A-1282074	Sequence 1282074, A

380	12	60.0	25	11	US-10-932-182A-89147	Sequence 89147, A	453	12	60.0	449	7	US-09-925-065A-500671	Sequence 500671,
c 381	12	60.0	25	11	US-10-932-182A-150829	Sequence 150829, A	454	12	60.0	451	7	US-09-925-065A-536155	Sequence 536155,
382	12	60.0	25	11	US-10-932-182A-89147	Sequence 89147, A	c 455	12	60.0	451	7	US-09-925-065A-555285	Sequence 555285,
c 383	12	60.0	25	11	US-10-932-182A-150829	Sequence 150829, A	c 456	12	60.0	451	7	US-09-925-065A-555286	Sequence 555286,
384	12	60.0	25	11	US-10-933-982-144592	Sequence 144592, A	457	12	60.0	452	7	US-09-925-065A-267197	Sequence 267197,
385	12	60.0	25	11	US-10-934-048A-48188	Sequence 48188, A	458	12	60.0	452	7	US-09-925-065A-267198	Sequence 267198,
c 386	12	60.0	25	11	US-10-934-048A-98450	Sequence 98450, A	459	12	60.0	452	12	US-10-301-480-344982	Sequence 344982,
c 387	12	60.0	25	11	US-11-121-849-1578	Sequence 1578, Ap	460	12	60.0	452	12	US-10-301-480-344983	Sequence 344983,
388	12	60.0	25	17	US-11-121-849-350134	Sequence 350134, A	461	12	60.0	452	12	US-10-301-480-958391	Sequence 958391,
389	12	60.0	25	17	US-11-121-849-369773	Sequence 369773, A	462	12	60.0	452	12	US-10-301-480-958392	Sequence 958392,
c 390	12	60.0	25	17	US-11-121-849-610569	Sequence 610569, A	c 463	12	60.0	455	7	US-09-925-065A-269825	Sequence 269825,
c 391	12	60.0	25	17	US-11-121-849-610570	Sequence 610570, A	c 464	12	60.0	457	7	US-09-925-065A-471320	Sequence 471320,
c 392	12	60.0	25	17	US-11-121-849-610571	Sequence 610571, A	c 465	12	60.0	457	7	US-09-925-065A-471321	Sequence 471321,
c 393	12	60.0	25	17	US-11-121-849-610572	Sequence 610572, A	c 466	12	60.0	457	7	US-09-925-065A-471322	Sequence 471322,
c 394	12	60.0	25	17	US-11-121-849-616319	Sequence 616319, A	c 467	12	60.0	457	12	US-10-301-480-390255	Sequence 390255,
c 395	12	60.0	25	17	US-11-136-527-34008	Sequence 34008, A	468	12	60.0	457	12	US-10-301-480-1003664	Sequence 1003664,
396	12	60.0	25	17	US-11-136-527-245621	Sequence 245621, A	469	12	60.0	457	12	US-09-925-065A-621852	Sequence 621852,
c 397	12	60.0	25	17	US-11-136-527-245633	Sequence 245633, A	470	12	60.0	468	7	US-09-925-065A-434070	Sequence 434070,
c 398	12	60.0	27	10	US-10-310-914A-136891	Sequence 136891, A	c 471	12	60.0	468	12	US-10-301-480-347367	Sequence 347367,
c 399	12	60.0	38	18	US-11-148-280-34	Sequence 34, Appl	c 472	12	60.0	468	12	US-10-301-480-960776	Sequence 960776,
400	12	60.0	40	18	US-11-286-819-2	Sequence 2, Appl	473	12	60.0	471	7	US-09-925-065A-146743	Sequence 146743,
401	12	60.0	50	17	US-11-175-859-111299	Sequence 111299, A	c 474	12	60.0	472	7	US-09-925-065A-222493	Sequence 222493,
c 402	12	60.0	70	10	US-10-310-914A-5287	Sequence 5287, Ap	475	12	60.0	472	18	US-11-096-568A-14159	Sequence 14159, A
c 403	12	60.0	74	10	US-10-310-914A-19140	Sequence 19140, A	c 476	12	60.0	472	18	US-11-251-226-1	Sequence 1, Appl
c 404	12	60.0	200	17	US-11-098-686-1553	Sequence 1553, Ap	477	12	60.0	475	7	US-09-925-065A-96552	Sequence 96552, A
c 405	12	60.0	200	17	US-11-098-686-7363	Sequence 7363, Ap	478	12	60.0	475	7	US-09-925-065A-96553	Sequence 96553, A
c 406	12	60.0	200	18	US-11-251-226-2	Sequence 2, Appl	479	12	60.0	475	7	US-09-925-065A-96554	Sequence 96554, A
c 407	12	60.0	201	10	US-10-995-561-28379	Sequence 28379, A	c 480	12	60.0	475	7	US-09-925-065A-504269	Sequence 504269,
c 408	12	60.0	201	10	US-10-995-561-28381	Sequence 28381, A	481	12	60.0	475	12	US-10-301-480-362593	Sequence 362593,
c 409	12	60.0	201	10	US-10-995-561-33001	Sequence 33001, A	482	12	60.0	475	12	US-10-301-480-976002	Sequence 976002,
c 410	12	60.0	201	10	US-10-995-561-33093	Sequence 33093, A	483	12	60.0	476	7	US-09-925-065A-765026	Sequence 765026,
c 411	12	60.0	201	10	US-10-995-561-33176	Sequence 33176, A	c 484	12	60.0	480	7	US-09-925-065A-578364	Sequence 578364,
c 412	12	60.0	201	10	US-10-995-561-41467	Sequence 41467, A	c 485	12	60.0	481	12	US-10-301-480-453112	Sequence 453112,
c 413	12	60.0	201	10	US-10-995-561-41468	Sequence 41468, A	c 486	12	60.0	481	12	US-10-301-480-1066521	Sequence 1066521,
c 414	12	60.0	201	10	US-10-995-561-47443	Sequence 47443, A	c 487	12	60.0	482	7	US-09-925-065A-384338	Sequence 384338,
c 415	12	60.0	201	10	US-10-995-561-48118	Sequence 48118, A	c 488	12	60.0	482	7	US-09-925-065A-580921	Sequence 580921,
c 416	12	60.0	201	10	US-10-995-561-55138	Sequence 55138, A	c 489	12	60.0	482	7	US-09-925-065A-580922	Sequence 580922,
c 417	12	60.0	201	10	US-10-995-561-57448	Sequence 57448, A	490	12	60.0	484	12	US-10-301-480-496258	Sequence 496258,
c 418	12	60.0	201	10	US-10-995-561-59078	Sequence 59078, A	491	12	60.0	484	12	US-10-301-480-1109667	Sequence 1109667,
c 419	12	60.0	201	10	US-10-995-561-67652	Sequence 67652, A	c 492	12	60.0	485	12	US-10-301-480-313888	Sequence 313888,
c 420	12	60.0	201	10	US-10-995-561-71828	Sequence 71828, A	c 493	12	60.0	485	12	US-10-301-480-927297	Sequence 927297,
c 421	12	60.0	201	17	US-11-124-368A-84141	Sequence 84141, Ap	494	12	60.0	487	7	US-09-925-065A-580966	Sequence 580966,
c 422	12	60.0	201	17	US-11-124-368A-11226	Sequence 11226, A	495	12	60.0	487	12	US-10-301-480-386396	Sequence 386396,
c 423	12	60.0	201	17	US-11-124-368A-13139	Sequence 13139, A	496	12	60.0	487	12	US-10-301-480-999805	Sequence 999805,
c 424	12	60.0	201	17	US-11-124-368A-13202	Sequence 13202, A	497	12	60.0	487	12	US-10-301-480-999805	Sequence 999805,
c 425	12	60.0	201	17	US-11-124-368A-18521	Sequence 18521, A	498	12	60.0	489	11	US-10-301-480-58630	Sequence 58630, A
c 426	12	60.0	201	17	US-11-124-367A-9388	Sequence 9388, Ap	499	12	60.0	489	7	US-10-301-480-672039	Sequence 672039,
c 427	12	60.0	201	17	US-11-124-367A-13760	Sequence 13760, A	500	12	60.0	490	7	US-09-925-065A-818693	Sequence 818693,
c 428	12	60.0	201	17	US-11-124-367A-17557	Sequence 17557, A	501	12	60.0	490	7	US-09-925-065A-818694	Sequence 818694,
c 429	12	60.0	201	17	US-11-124-367A-19654	Sequence 19654, A	c 502	12	60.0	491	7	US-09-925-065A-230231	Sequence 230231,
c 430	12	60.0	201	17	US-11-124-367A-19948	Sequence 19948, A	c 503	12	60.0	494	7	US-09-925-065A-310969	Sequence 310969,
c 431	12	60.0	201	17	US-11-124-367A-19949	Sequence 19949, A	504	12	60.0	495	7	US-09-925-065A-731624	Sequence 731624,
c 432	12	60.0	201	17	US-11-124-367A-20030	Sequence 20030, A	505	12	60.0	495	7	US-09-925-065A-381138	Sequence 381138,
c 433	12	60.0	201	17	US-11-124-367A-20102	Sequence 20102, A	506	12	60.0	495	7	US-09-925-065A-381139	Sequence 381139,
c 434	12	60.0	201	17	US-11-124-367A-32739	Sequence 32739, A	c 507	12	60.0	495	7	US-09-925-065A-625837	Sequence 625837,
c 435	12	60.0	201	17	US-11-124-367A-32740	Sequence 32740, A	c 508	12	60.0	499	7	US-09-925-065A-731731	Sequence 731731,
c 436	12	60.0	351	7	US-09-925-065A-596207	Sequence 596207, A	c 509	12	60.0	499	7	US-09-925-065A-314968	Sequence 314968,
c 437	12	60.0	357	7	US-09-925-065A-16043	Sequence 16043, A	c 510	12	60.0	499	11	US-10-301-480-71481	Sequence 71481, A
c 438	12	60.0	357	11	US-10-301-480-117280	Sequence 117280, A	c 511	12	60.0	499	11	US-10-301-480-71482	Sequence 71482, A
c 439	12	60.0	357	12	US-10-301-480-730689	Sequence 730689, A	c 512	12	60.0	499	12	US-10-301-480-684890	Sequence 684890,
c 440	12	60.0	363	11	US-10-932-182A-5522	Sequence 5522, Ap	c 513	12	60.0	499	12	US-10-301-480-684891	Sequence 684891,
c 441	12	60.0	363	11	US-10-932-182A-5522	Sequence 5522, Ap	514	12	60.0	500	12	US-10-301-480-450237	Sequence 450237,
c 442	12	60.0	368	11	US-10-301-480-11386	Sequence 11386, A	515	12	60.0	500	12	US-10-301-480-450238	Sequence 450238,
c 443	12	60.0	408	12	US-10-301-480-624795	Sequence 624795, A	516	12	60.0	500	12	US-10-301-480-1063646	Sequence 1063646,
c 444	12	60.0	422	7	US-09-925-065A-555287	Sequence 555287, A	517	12	60.0	500	12	US-10-301-480-1063647	Sequence 1063647,
c 445	12	60.0	430	12	US-10-301-480-458874	Sequence 458874, A	518	12	60.0	501	7	US-09-925-065A-590833	Sequence 590833,
c 446	12	60.0	430	12	US-10-301-480-1072283	Sequence 1072283, A	c 519	12	60.0	501	7	US-09-925-065A-590834	Sequence 590834,
c 447	12	60.0	432	7	US-09-925-065A-390569	Sequence 390569, A	c 520	12	60.0	501	17	US-11-000-688-1063	Sequence 1063, Ap
c 448	12	60.0	437	7	US-09-925-065A-315234	Sequence 315234, A	521	12	60.0	505	7	US-09-925-065A-548845	Sequence 548845,
c 449	12	60.0	442	11	US-10-301-480-1744	Sequence 1744, Ap	522	12	60.0	505	12	US-10-301-480-526795	Sequence 526795,
c 450	12	60.0	442	12	US-10-301-480-615153	Sequence 615153, A	523	12	60.0	505	12	US-10-301-480-604921	Sequence 604921,
c 451	12	60.0	446	7	US-09-925-065A-468946	Sequence 468946, A	524	12	60.0	505	12	US-10-301-480-1140204	Sequence 1140204,
c 452	12	60.0	449	7	US-09-925-065A-500670	Sequence 500670, A	525	12	60.0	505	12	US-10-301-480-1218330	Sequence 1218330,

C 526	12	60.0	507	7	US-09-925-065A-659051	Sequence 659051, A	599	12	60.0	526	7	US-09-925-065A-802099	Sequence 802099, A
527	12	60.0	508	7	US-09-925-065A-338703	Sequence 338703, A	600	12	60.0	526	7	US-09-925-065A-855818	Sequence 855818, A
528	12	60.0	508	7	US-09-925-065A-338704	Sequence 338704, A	C 601	12	60.0	526	10	US-10-454-437-121	Sequence 121, App
C 529	12	60.0	510	7	US-09-925-065A-257180	Sequence 257180, A	C 602	12	60.0	527	7	US-09-925-065A-20317	Sequence 20317, A
C 530	12	60.0	510	7	US-09-925-065A-257181	Sequence 257181, A	C 603	12	60.0	527	11	US-10-301-480-121554	Sequence 121554, A
C 531	12	60.0	510	7	US-09-925-065A-295515	Sequence 295515, A	C 604	12	60.0	527	12	US-10-301-480-734963	Sequence 734963, A
C 532	12	60.0	510	7	US-09-925-065A-579969	Sequence 579969, A	C 605	12	60.0	528	7	US-09-925-065A-242976	Sequence 242976, A
533	12	60.0	512	7	US-09-925-065A-388359	Sequence 388359, A	C 606	12	60.0	528	7	US-09-925-065A-242977	Sequence 242977, A
534	12	60.0	512	12	US-10-301-480-345209	Sequence 345209, A	C 607	12	60.0	529	7	US-09-925-065A-135396	Sequence 135396, A
535	12	60.0	512	12	US-10-301-480-345210	Sequence 345210, A	C 608	12	60.0	529	12	US-10-301-480-347364	Sequence 347364, A
536	12	60.0	512	12	US-10-301-480-958618	Sequence 958618, A	C 609	12	60.0	529	12	US-09-925-065A-960773	Sequence 960773, A
537	12	60.0	512	12	US-10-301-480-958619	Sequence 958619, A	C 610	12	60.0	530	7	US-09-925-065A-180588	Sequence 180588, A
538	12	60.0	514	7	US-09-925-065A-331814	Sequence 331814, A	C 611	12	60.0	530	7	US-09-925-065A-269822	Sequence 269822, A
539	12	60.0	514	7	US-09-925-065A-331815	Sequence 331815, A	C 612	12	60.0	530	7	US-09-925-065A-641775	Sequence 641775, A
540	12	60.0	514	7	US-09-925-065A-331816	Sequence 331816, A	C 613	12	60.0	530	7	US-09-925-065A-641776	Sequence 641776, A
541	12	60.0	516	7	US-09-925-065A-267439	Sequence 267439, A	C 614	12	60.0	531	7	US-09-925-065A-3628	Sequence 3628, App
542	12	60.0	516	7	US-09-925-065A-267438	Sequence 267438, A	C 615	12	60.0	531	7	US-09-925-065A-280596	Sequence 280596, A
C 543	12	60.0	516	12	US-10-301-480-335865	Sequence 335865, A	C 616	12	60.0	531	11	US-10-301-480-104865	Sequence 104865, A
C 544	12	60.0	516	12	US-10-301-480-335866	Sequence 335866, A	C 617	12	60.0	531	12	US-10-301-480-104865	Sequence 104865, A
C 545	12	60.0	516	12	US-10-301-480-405575	Sequence 405575, A	C 618	12	60.0	532	7	US-09-925-065A-20376	Sequence 20376, A
C 546	12	60.0	516	12	US-10-301-480-405576	Sequence 405576, A	C 619	12	60.0	532	7	US-09-925-065A-558646	Sequence 558646, A
547	12	60.0	516	12	US-10-301-480-405577	Sequence 405577, A	C 620	12	60.0	532	11	US-10-301-480-121613	Sequence 121613, A
C 548	12	60.0	516	12	US-10-301-480-949274	Sequence 949274, A	C 621	12	60.0	532	12	US-10-301-480-735022	Sequence 735022, A
C 549	12	60.0	516	12	US-10-301-480-949275	Sequence 949275, A	C 622	12	60.0	533	7	US-09-925-065A-17253	Sequence 17253, A
550	12	60.0	516	12	US-10-301-480-1018984	Sequence 1018984, A	C 623	12	60.0	533	7	US-09-925-065A-136502	Sequence 136502, A
551	12	60.0	516	12	US-10-301-480-1018985	Sequence 1018985, A	C 624	12	60.0	533	7	US-09-925-065A-256927	Sequence 256927, A
552	12	60.0	516	12	US-10-301-480-1018986	Sequence 1018986, A	C 625	12	60.0	533	7	US-09-925-065A-256928	Sequence 256928, A
553	12	60.0	518	7	US-09-925-065A-592375	Sequence 592375, A	C 626	12	60.0	533	7	US-09-925-065A-256929	Sequence 256929, A
554	12	60.0	518	7	US-09-925-065A-899092	Sequence 899092, A	C 627	12	60.0	533	11	US-10-301-480-118490	Sequence 118490, A
C 555	12	60.0	518	12	US-10-301-480-602370	Sequence 602370, A	C 628	12	60.0	533	12	US-10-301-480-731899	Sequence 731899, A
C 556	12	60.0	518	12	US-10-301-480-1215779	Sequence 1215779, A	C 629	12	60.0	533	12	US-10-301-480-731899	Sequence 731899, A
C 557	12	60.0	519	7	US-09-925-065A-801150	Sequence 801150, A	C 630	12	60.0	533	12	US-10-301-480-1135974	Sequence 1135974, A
C 558	12	60.0	519	7	US-09-925-065A-854985	Sequence 854985, A	C 631	12	60.0	534	12	US-10-301-480-601071	Sequence 601071, A
C 559	12	60.0	519	12	US-10-301-480-390003	Sequence 390003, A	C 632	12	60.0	534	12	US-10-301-480-1214480	Sequence 1214480, A
C 560	12	60.0	519	12	US-10-301-480-390004	Sequence 390004, A	C 633	12	60.0	535	7	US-09-925-065A-20412	Sequence 20412, A
C 561	12	60.0	519	12	US-10-301-480-1003412	Sequence 1003412, A	C 634	12	60.0	535	7	US-09-925-065A-466960	Sequence 466960, A
C 562	12	60.0	519	12	US-10-301-480-1003413	Sequence 1003413, A	C 635	12	60.0	535	11	US-10-301-480-23356	Sequence 23356, A
C 563	12	60.0	520	12	US-10-301-480-466104	Sequence 466104, A	C 636	12	60.0	535	11	US-10-301-480-121649	Sequence 121649, A
564	12	60.0	520	12	US-10-301-480-466105	Sequence 466105, A	C 637	12	60.0	535	12	US-10-301-480-636765	Sequence 636765, A
565	12	60.0	520	12	US-10-301-480-1079513	Sequence 1079513, A	C 638	12	60.0	535	12	US-10-301-480-735058	Sequence 735058, A
566	12	60.0	520	12	US-10-301-480-1079514	Sequence 1079514, A	C 639	12	60.0	537	11	US-10-301-480-20161	Sequence 20161, A
567	12	60.0	522	7	US-09-925-065A-58270	Sequence 58270, A	C 640	12	60.0	537	12	US-10-301-480-63570	Sequence 63570, A
568	12	60.0	522	11	US-10-301-480-159508	Sequence 159508, A	C 641	12	60.0	538	7	US-09-925-065A-467647	Sequence 467647, A
569	12	60.0	522	12	US-10-301-480-456789	Sequence 456789, A	C 642	12	60.0	538	12	US-10-301-480-232039	Sequence 232039, A
570	12	60.0	522	12	US-10-301-480-772917	Sequence 772917, A	C 643	12	60.0	538	12	US-10-301-480-271329	Sequence 271329, A
571	12	60.0	522	12	US-10-301-480-1070198	Sequence 1070198, A	C 644	12	60.0	538	12	US-10-301-480-323867	Sequence 323867, A
572	12	60.0	523	7	US-09-925-065A-36328	Sequence 36328, A	C 645	12	60.0	538	12	US-10-301-480-323867	Sequence 323867, A
573	12	60.0	523	7	US-09-925-065A-36329	Sequence 36329, A	C 646	12	60.0	538	12	US-10-301-480-824548	Sequence 824548, A
C 574	12	60.0	523	7	US-09-925-065A-48625	Sequence 48625, A	C 647	12	60.0	538	12	US-10-301-480-824548	Sequence 824548, A
C 575	12	60.0	523	7	US-09-925-065A-48626	Sequence 48626, A	C 648	12	60.0	538	12	US-10-301-480-824548	Sequence 824548, A
C 576	12	60.0	523	7	US-09-925-065A-533378	Sequence 533378, A	C 649	12	60.0	538	12	US-10-301-480-937276	Sequence 937276, A
C 577	12	60.0	523	7	US-09-925-065A-533379	Sequence 533379, A	C 650	12	60.0	538	12	US-10-301-480-937277	Sequence 937277, A
578	12	60.0	523	11	US-10-301-480-137566	Sequence 137566, A	C 651	12	60.0	540	12	US-10-301-480-483642	Sequence 483642, A
579	12	60.0	523	11	US-10-301-480-137567	Sequence 137567, A	C 652	12	60.0	540	12	US-10-301-480-1097051	Sequence 1097051, A
C 580	12	60.0	523	11	US-10-301-480-149863	Sequence 149863, A	C 653	12	60.0	541	12	US-09-925-065A-479278	Sequence 479278, A
C 581	12	60.0	523	11	US-10-301-480-149864	Sequence 149864, A	C 654	12	60.0	541	12	US-10-301-480-484529	Sequence 484529, A
582	12	60.0	523	12	US-10-301-480-357544	Sequence 357544, A	C 655	12	60.0	541	12	US-10-301-480-484530	Sequence 484530, A
583	12	60.0	523	12	US-10-301-480-549169	Sequence 549169, A	C 656	12	60.0	541	12	US-10-301-480-484531	Sequence 484531, A
584	12	60.0	523	12	US-10-301-480-750975	Sequence 750975, A	C 657	12	60.0	541	12	US-10-301-480-523186	Sequence 523186, A
585	12	60.0	523	12	US-10-301-480-750976	Sequence 750976, A	C 658	12	60.0	541	12	US-10-301-480-1098038	Sequence 1098038, A
C 586	12	60.0	523	12	US-10-301-480-763272	Sequence 763272, A	C 659	12	60.0	541	12	US-10-301-480-1098039	Sequence 1098039, A
C 587	12	60.0	523	12	US-10-301-480-763273	Sequence 763273, A	C 660	12	60.0	541	12	US-10-301-480-1098040	Sequence 1098040, A
588	12	60.0	523	12	US-10-301-480-970953	Sequence 970953, A	C 661	12	60.0	542	7	US-09-925-065A-1136595	Sequence 1136595, A
589	12	60.0	523	12	US-10-301-480-1162578	Sequence 1162578, A	C 662	12	60.0	542	7	US-09-925-065A-18365	Sequence 18365, A
C 590	12	60.0	525	7	US-09-925-065A-24373	Sequence 24373, A	C 663	12	60.0	542	7	US-09-925-065A-596333	Sequence 596333, A
591	12	60.0	525	7	US-09-925-065A-24374	Sequence 24374, A	C 664	12	60.0	543	7	US-09-925-065A-313019	Sequence 313019, A
592	12	60.0	525	11	US-10-301-480-125610	Sequence 125610, A	C 665	12	60.0	543	7	US-09-925-065A-662714	Sequence 662714, A
593	12	60.0	525	11	US-10-301-480-125611	Sequence 125611, A	C 666	12	60.0	543	7	US-09-925-065A-860019	Sequence 860019, A
C 594	12	60.0	525	12	US-10-301-480-372374	Sequence 372374, A	C 667	12	60.0	543	7	US-09-925-065A-860020	Sequence 860020, A
595	12	60.0	525	12	US-10-301-480-739019	Sequence 739019, A	C 668	12	60.0	546	7	US-09-925-065A-824296	Sequence 824296, A
596	12	60.0	525	12	US-10-301-480-739020	Sequence 739020, A	C 669	12	60.0	547	7	US-09-925-065A-295077	Sequence 295077, A
C 597	12	60.0	525	12	US-10-301-480-985783	Sequence 985783, A	C 670	12	60.0	548	7	US-09-925-065A-467286	Sequence 467286, A
598	12	60.0	526	7	US-09-925-065A-610293	Sequence 610293, A	C 671	12	60.0	548	7	US-09-925-065A-617715	Sequence 617715, A

c 672	12	60.0	548	7	US-09-925-065A-617716	Sequence 617716,	745	12	60.0	562	7	US-09-925-065A-140394	Sequence 140394,
c 673	12	60.0	549	7	US-09-925-065A-139724	Sequence 139724,	746	12	60.0	562	7	US-09-925-065A-325431	Sequence 325431,
c 674	12	60.0	549	7	US-09-925-065A-596330	Sequence 596330,	747	12	60.0	562	7	US-09-925-065A-358318	Sequence 358318,
c 675	12	60.0	549	7	US-09-925-065A-596331	Sequence 596331,	748	12	60.0	562	7	US-09-925-065A-430081	Sequence 430081,
c 676	12	60.0	549	7	US-09-925-065A-596332	Sequence 596332,	c 749	12	60.0	562	7	US-09-925-065A-882660	Sequence 882660,
c 677	12	60.0	549	7	US-09-925-065A-596333	Sequence 596333,	c 750	12	60.0	562	7	US-09-925-065A-882661	Sequence 882661,
c 678	12	60.0	549	12	US-10-301-480-437331	Sequence 437331,	c 751	12	60.0	562	7	US-09-925-065A-109713	Sequence 109713,
c 679	12	60.0	549	12	US-10-301-480-1050740	Sequence 1050740,	c 752	12	60.0	562	7	US-09-925-065A-109332	Sequence 109332,
c 680	12	60.0	550	7	US-09-925-065A-634016	Sequence 634016,	c 753	12	60.0	563	7	US-09-925-065A-112206	Sequence 112206,
c 681	12	60.0	550	12	US-10-301-480-335619	Sequence 335619,	c 754	12	60.0	563	7	US-09-925-065A-751576	Sequence 751576,
c 682	12	60.0	550	12	US-10-301-480-335620	Sequence 335620,	c 755	12	60.0	563	7	US-09-925-065A-751577	Sequence 751577,
c 683	12	60.0	550	12	US-10-301-480-335621	Sequence 335621,	756	12	60.0	563	12	US-10-301-480-492913	Sequence 492913,
c 684	12	60.0	550	12	US-10-301-480-366942	Sequence 366942,	757	12	60.0	563	12	US-10-301-480-1106322	Sequence 1106322,
c 685	12	60.0	550	12	US-10-301-480-366942	Sequence 366942,	c 758	12	60.0	564	7	US-09-925-065A-22433	Sequence 22433, A
c 686	12	60.0	550	12	US-10-301-480-949028	Sequence 949028,	c 759	12	60.0	564	7	US-09-925-065A-329249	Sequence 329249, A
c 687	12	60.0	550	12	US-10-301-480-949029	Sequence 949029,	760	12	60.0	564	7	US-09-925-065A-329250	Sequence 329250,
c 688	12	60.0	550	12	US-10-301-480-949030	Sequence 949030,	c 761	12	60.0	564	11	US-10-301-480-123670	Sequence 123670,
c 689	12	60.0	551	7	US-09-925-065A-980351	Sequence 980351,	c 762	12	60.0	564	12	US-10-301-480-737079	Sequence 737079,
c 690	12	60.0	551	11	US-10-301-480-113437	Sequence 113437,	c 763	12	60.0	565	7	US-09-925-065A-274416	Sequence 274416,
c 691	12	60.0	551	12	US-10-301-480-113437	Sequence 113437,	c 764	12	60.0	565	7	US-09-925-065A-274417	Sequence 274417,
c 692	12	60.0	551	12	US-10-301-480-234784	Sequence 234784,	c 765	12	60.0	565	7	US-09-925-065A-274418	Sequence 274418,
c 693	12	60.0	551	12	US-10-301-480-403147	Sequence 403147,	c 766	12	60.0	565	7	US-09-925-065A-274418	Sequence 274418,
c 694	12	60.0	551	12	US-10-301-480-403148	Sequence 403148,	c 767	12	60.0	565	7	US-09-925-065A-398055	Sequence 398055,
c 695	12	60.0	551	12	US-10-301-480-726846	Sequence 726846,	c 768	12	60.0	565	7	US-09-925-065A-398056	Sequence 398056,
c 696	12	60.0	551	12	US-10-301-480-848193	Sequence 848193,	c 769	12	60.0	565	11	US-10-301-480-208994	Sequence 208994,
c 697	12	60.0	551	12	US-10-301-480-1016556	Sequence 1016556,	c 770	12	60.0	565	12	US-10-301-480-327291	Sequence 327291,
c 698	12	60.0	551	12	US-10-301-480-1016557	Sequence 1016557,	c 771	12	60.0	565	12	US-10-301-480-822403	Sequence 822403,
c 699	12	60.0	552	7	US-09-925-065A-266330	Sequence 266330, A	c 772	12	60.0	565	12	US-10-301-480-940700	Sequence 940700,
c 700	12	60.0	552	7	US-09-925-065A-266331	Sequence 266331, A	c 773	12	60.0	566	7	US-09-925-065A-325432	Sequence 325432,
c 701	12	60.0	552	7	US-09-925-065A-266332	Sequence 266332, A	c 774	12	60.0	566	11	US-10-301-480-14748	Sequence 14748, A
c 702	12	60.0	552	7	US-09-925-065A-588233	Sequence 588233, A	c 775	12	60.0	566	12	US-10-301-480-628157	Sequence 628157,
c 703	12	60.0	552	7	US-09-925-065A-605092	Sequence 605092, A	c 776	12	60.0	567	7	US-09-925-065A-805321	Sequence 805321,
c 704	12	60.0	552	11	US-10-301-480-127867	Sequence 127867,	c 777	12	60.0	567	7	US-09-925-065A-805322	Sequence 805322,
c 705	12	60.0	552	11	US-10-301-480-127868	Sequence 127868,	c 778	12	60.0	567	11	US-10-301-480-61533	Sequence 61533, A
c 706	12	60.0	552	11	US-10-301-480-127869	Sequence 127869,	c 779	12	60.0	568	12	US-10-301-480-674942	Sequence 674942,
c 707	12	60.0	552	12	US-10-301-480-741276	Sequence 741276,	c 780	12	60.0	568	12	US-10-301-480-354771	Sequence 354771,
c 708	12	60.0	552	12	US-10-301-480-741277	Sequence 741277,	c 781	12	60.0	568	12	US-10-301-480-522861	Sequence 522861,
c 709	12	60.0	553	7	US-09-925-065A-212789	Sequence 212789,	c 782	12	60.0	568	12	US-10-301-480-968180	Sequence 968180,
c 710	12	60.0	553	7	US-09-925-065A-212790	Sequence 212790,	c 783	12	60.0	569	12	US-10-301-480-1136270	Sequence 1136270,
c 711	12	60.0	553	7	US-09-925-065A-696692	Sequence 696692,	c 784	12	60.0	569	7	US-09-925-065A-239752	Sequence 239752,
c 712	12	60.0	553	7	US-09-925-065A-696692	Sequence 696692,	c 785	12	60.0	569	7	US-09-925-065A-247168	Sequence 247168,
c 713	12	60.0	554	11	US-10-301-480-891752	Sequence 891752,	c 786	12	60.0	569	12	US-10-301-480-284270	Sequence 284270,
c 714	12	60.0	554	12	US-10-301-480-71211	Sequence 71211, A	c 787	12	60.0	569	12	US-10-301-480-429577	Sequence 429577,
c 715	12	60.0	554	12	US-10-301-480-321544	Sequence 321544,	c 788	12	60.0	569	12	US-10-301-480-450625	Sequence 450625,
c 716	12	60.0	554	12	US-10-301-480-591381	Sequence 591381,	c 789	12	60.0	569	12	US-10-301-480-450626	Sequence 450626,
c 717	12	60.0	554	12	US-10-301-480-684620	Sequence 684620,	c 790	12	60.0	569	12	US-10-301-480-450627	Sequence 450627,
c 718	12	60.0	554	12	US-10-301-480-934953	Sequence 934953,	c 791	12	60.0	569	12	US-10-301-480-897679	Sequence 897679,
c 719	12	60.0	554	12	US-10-301-480-1204790	Sequence 1204790,	c 792	12	60.0	569	12	US-10-301-480-1042986	Sequence 1042986,
c 720	12	60.0	555	7	US-09-925-065A-419516	Sequence 419516,	c 793	12	60.0	569	12	US-10-301-480-1064034	Sequence 1064034,
c 721	12	60.0	555	7	US-09-925-065A-419517	Sequence 419517,	c 794	12	60.0	569	12	US-10-301-480-1064035	Sequence 1064035,
c 722	12	60.0	555	7	US-09-925-065A-419518	Sequence 419518,	c 795	12	60.0	569	12	US-10-301-480-1064036	Sequence 1064036,
c 723	12	60.0	555	11	US-10-301-480-48548	Sequence 48548, A	c 796	12	60.0	571	7	US-09-925-065A-97379	Sequence 97379, A
c 724	12	60.0	555	12	US-10-301-480-661957	Sequence 661957,	c 797	12	60.0	571	7	US-09-925-065A-97380	Sequence 97380, A
c 725	12	60.0	556	7	US-09-925-065A-290029	Sequence 290029,	c 798	12	60.0	571	7	US-09-925-065A-97381	Sequence 97381, A
c 726	12	60.0	556	7	US-09-925-065A-366960	Sequence 366960,	c 799	12	60.0	572	7	US-09-925-065A-295820	Sequence 295820,
c 727	12	60.0	556	11	US-10-301-480-10847	Sequence 10847, A	c 800	12	60.0	572	7	US-09-925-065A-310968	Sequence 310968,
c 728	12	60.0	556	12	US-10-301-480-112396	Sequence 112396,	c 801	12	60.0	572	7	US-09-925-065A-310970	Sequence 310970,
c 729	12	60.0	556	12	US-10-301-480-624256	Sequence 624256,	c 802	12	60.0	572	11	US-09-925-065A-381911	Sequence 381911,
c 730	12	60.0	556	12	US-10-301-480-925805	Sequence 925805,	c 803	12	60.0	572	11	US-10-301-480-198317	Sequence 198317,
c 731	12	60.0	558	12	US-10-301-480-228750	Sequence 228750,	c 804	12	60.0	572	11	US-10-301-480-198318	Sequence 198318,
c 732	12	60.0	558	12	US-10-301-480-371967	Sequence 371967,	c 805	12	60.0	572	11	US-10-301-480-198319	Sequence 198319,
c 733	12	60.0	558	12	US-10-301-480-985376	Sequence 985376,	c 806	12	60.0	572	12	US-10-301-480-811726	Sequence 811726,
c 734	12	60.0	559	7	US-09-925-065A-195050	Sequence 195050,	c 807	12	60.0	572	12	US-10-301-480-811727	Sequence 811727,
c 735	12	60.0	559	7	US-09-925-065A-366475	Sequence 366475,	c 808	12	60.0	572	12	US-10-301-480-811728	Sequence 811728,
c 736	12	60.0	559	11	US-10-301-480-71210	Sequence 71210, A	c 809	12	60.0	573	7	US-09-925-065A-30968	Sequence 30968, A
c 737	12	60.0	559	11	US-10-301-480-71212	Sequence 71212, A	c 810	12	60.0	573	7	US-09-925-065A-412244	Sequence 412244,
c 738	12	60.0	559	12	US-10-301-480-436905	Sequence 436905,	c 811	12	60.0	573	11	US-10-301-480-100482	Sequence 100482,
c 739	12	60.0	559	12	US-10-301-480-684619	Sequence 684619,	c 812	12	60.0	573	11	US-10-301-480-132206	Sequence 132206,
c 740	12	60.0	559	12	US-10-301-480-684621	Sequence 684621,	c 813	12	60.0	573	11	US-10-301-480-235391	Sequence 235391,
c 741	12	60.0	559	12	US-10-301-480-684621	Sequence 684621,	c 814	12	60.0	573	12	US-10-301-480-235392	Sequence 235392,
c 742	12	60.0	560	7	US-09-925-065A-1050314	Sequence 1050314,	c 815	12	60.0	573	12	US-10-301-480-235393	Sequence 235393,
c 743	12	60.0	562	7	US-09-925-065A-228388	Sequence 228388,	c 816	12	60.0	573	12	US-10-301-480-713891	Sequence 713891,
c 744	12	60.0	562	7	US-09-925-065A-124637	Sequence 124637,	c 817	12	60.0	573	12	US-10-301-480-713891	Sequence 713891,
	12	60.0	562	7	US-09-925-065A-140391	Sequence 140391,		12	60.0	573	12	US-10-301-480-745615	Sequence 745615,
	12	60.0	562	7	US-09-925-065A-140392	Sequence 140392,		12	60.0	573	12	US-10-301-480-848800	Sequence 848800,



[illegible]

```
c 964      12 60.0 593 7 US-09-925-065A-311157
c 965      12 60.0 593 7 US-09-925-065A-311158
c 966      12 60.0 593 7 US-09-925-065A-311159
c 967      12 60.0 593 7 US-09-925-065A-311160
c 968      12 60.0 593 7 US-09-925-065A-311160
c 968      12 60.0 593 7 US-09-925-065A-771763
c 969      12 60.0 594 7 US-09-925-065A-455630
c 970      12 60.0 594 7 US-09-925-065A-455631
c 971      12 60.0 594 12 US-10-301-480-386395
c 972      12 60.0 594 12 US-10-301-480-386397
c 973      12 60.0 594 12 US-10-301-480-999804
c 974      12 60.0 594 12 US-10-301-480-999806
c 975      12 60.0 595 7 US-09-925-065A-155954
c 976      12 60.0 595 7 US-09-925-065A-155955
c 977      12 60.0 595 7 US-09-925-065A-553991
c 978      12 60.0 596 7 US-09-925-065A-746858
c 979      12 60.0 597 7 US-09-925-065A-805324
c 980      12 60.0 597 7 US-09-925-065A-805325
c 981      12 60.0 598 7 US-09-925-065A-165416
c 982      12 60.0 598 7 US-09-925-065A-257179
c 983      12 60.0 598 7 US-09-925-065A-480849
c 984      12 60.0 598 7 US-09-925-065A-634015
c 985      12 60.0 598 7 US-09-925-065A-934404
c 986      12 60.0 598 10 US-10-750-185-1602
c 987      12 60.0 598 10 US-10-750-623-1602
c 988      12 60.0 598 12 US-10-301-480-249300
c 989      12 60.0 598 12 US-10-301-480-249301
c 990      12 60.0 598 12 US-10-301-480-586925
c 991      12 60.0 598 12 US-10-301-480-862709
c 992      12 60.0 598 12 US-10-301-480-862710
c 993      12 60.0 598 12 US-10-301-480-1200334
c 994      12 60.0 599 11 US-10-301-480-199828
c 995      12 60.0 599 12 US-10-301-480-813237
c 996      12 60.0 600 7 US-09-925-065A-404717
c 997      12 60.0 600 7 US-09-925-065A-404718
c 998      12 60.0 600 7 US-09-925-065A-404719
c 999      12 60.0 600 7 US-09-925-065A-786681
c1000     12 60.0 600 10 US-10-750-185-3788
```

## ALIGNMENTS

```
RESULT 1
US-11-136-527-226522/c
; Sequence 226522, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 226522
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Probe
US-11-136-527-226522
Query Match 100.0%; Score 20; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GCCTTGGTTGGCAACACAT 20
        |||||
Db      21 GCCTTGGTTGGCAACACAT 2
```

```
Sequence 311157,
Sequence 311158,
Sequence 311159,
Sequence 311160,
Sequence 771763,
Sequence 455630,
Sequence 455631,
Sequence 386395,
Sequence 386397,
Sequence 999804,
Sequence 999806,
Sequence 155954,
Sequence 155955,
Sequence 553991,
Sequence 746858,
Sequence 805324,
Sequence 805325,
Sequence 165416,
Sequence 257179,
Sequence 480849,
Sequence 634015,
Sequence 934404,
Sequence 1602, Ap
Sequence 1602, Ap
Sequence 249300,
Sequence 249301,
Sequence 586925,
Sequence 862709,
Sequence 862710,
Sequence 1200334,
Sequence 199828,
Sequence 813237,
Sequence 404717,
Sequence 404718,
Sequence 404719,
Sequence 786681,
Sequence 3788, Ap
```

## RESULT 2

```
US-10-920-152-12/c
; Sequence 12, Application US/10920152
; Publication No. US20060057584A1
; GENERAL INFORMATION:
; APPLICANT: Baban, Soheyl
; APPLICANT: Cherry, Elana
; APPLICANT: Gosselin, Diane
; APPLICANT: Hugo, Patrice
; APPLICANT: Malette, Brigitte
; APPLICANT: Miron, Pierre
; APPLICANT: Prive, Charles
; APPLICANT: Shazand, Kamran
; TITLE OF INVENTION: ENDOMETRIOSIS-RELATED MARKERS AND USES THEREOF
; FILE REFERENCE: 5600.71
; CURRENT APPLICATION NUMBER: US/10/920,152
; CURRENT FILING DATE: 2004-08-17
; PRIOR APPLICATION NUMBER: US/09/794,928
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 60/225,745
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 60/185,063
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 254
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-920-152-12
```

```
Query Match 100.0%; Score 20; DB 12; Length 254;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GCCTTGGTTGGCAACACAT 20
        |||||
Db      90 GCCTTGGTTGGCAACACAT 71
```

## RESULT 3

```
US-11-136-527-4790/c
; Sequence 4790, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4790
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4790
```

```
Query Match 100.0%; Score 20; DB 17; Length 600;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GCCTTGGTTGGCAACACAT 20
        |||||
Db      428 GCCTTGGTTGGCAACACAT 409
```

## RESULT 4



US-11-136-527-694/c  
; Sequence 694, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-041000 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; CURRENT FILING DATE: 2005-05-25  
; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 694  
; LENGTH: 1803  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
US-11-136-527-694

Query Match 100.0%; Score 20; DB 17; Length 1803;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 1631 GCTTTGGTTGGCAACACAT 1612

RESULT 5  
US-10-469-469-192/c  
; Sequence 192, Application US/10469469  
; Publication No. US20060079493A1  
; GENERAL INFORMATION:  
; APPLICANT: FRITZ, LAWRENCE C.  
; APPLICANT: BURROWS, FRANCIS J.  
; TITLE OF INVENTION: METHODS FOR TREATING GENETICALLY-DEFINED PROLIFERATIVE  
; TITLE OF INVENTION: DISORDERS WITH HSP90 INHIBITORS  
; FILE REFERENCE: CON-0010-USN  
; CURRENT APPLICATION NUMBER: US/10/469,469  
; CURRENT FILING DATE: 2003-08-27  
; PRIOR APPLICATION NUMBER: PCT/US02/06518  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/272,751  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 192  
; LENGTH: 3517  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-469-469-192

Query Match 100.0%; Score 20; DB 13; Length 3517;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACAT 20  
|||||  
Db 2022 GCTTTGGTTGGCAACACAT 2003

RESULT 6  
US-11-136-527-34022/c  
; Sequence 34022, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-041000 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; CURRENT FILING DATE: 2005-05-25

; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 34022  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Probe  
US-11-136-527-34022

Query Match 95.0%; Score 19; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.57;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTTGGTTGGCAACACAT 20  
|||||  
Db 25 CTTTGGTTGGCAACACAT 7

RESULT 7  
US-11-136-527-226492/c  
; Sequence 226492, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-041000 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; CURRENT FILING DATE: 2005-05-25  
; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 226492  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Probe  
US-11-136-527-226492

Query Match 95.0%; Score 19; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.57;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTTGGTTGGCAACACAT 20  
|||||  
Db 25 CTTTGGTTGGCAACACAT 7

RESULT 8  
US-11-136-527-226494/c  
; Sequence 226494, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-041000 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; CURRENT FILING DATE: 2005-05-25  
; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 226494  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:

; OTHER INFORMATION: Probe  
US-11-136-527-226494

Query Match 95.0%; Score 19; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.57;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGCAACACA 19  
|||||  
Db 19 GCTTTGGTTGGCAACACA 1

## RESULT 9

US-11-136-527-34046/c  
; Sequence 34046, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-041000 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; CURRENT FILING DATE: 2005-05-25  
; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 34046  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Probe  
US-11-136-527-34046

Query Match 90.0%; Score 18; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTTGGTTGGCAACACAT 20  
|||||  
Db 25 TTTGGTTGGCAACACAT 8

## RESULT 10

US-11-136-527-34023/c  
; Sequence 34023, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-041000 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; CURRENT FILING DATE: 2005-05-25  
; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 34023  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Probe  
US-11-136-527-34023

Query Match 85.0%; Score 17; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TTTGGTTGGCAACACAT 20  
|||||

Db 25 TTGGTTGGCAACACAT 9

## RESULT 11

US-11-136-527-34042/c  
; Sequence 34042, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-041000 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; CURRENT FILING DATE: 2005-05-25  
; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 34042  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Probe  
US-11-136-527-34042

Query Match 80.0%; Score 16; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TGGTTGGCAACACAT 20  
|||||  
Db 25 TGGTTGGCAACACAT 10

## RESULT 12

US-10-301-480-546810  
; Sequence 546810, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 546810  
; LENGTH: 995  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-546810

Query Match 80.0%; Score 16; DB 12; Length 995;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TTGGTTGGCAACACA 19  
|||||  
Db 823 TTGGTTGGCAACACA 838

## RESULT 13

US-10-301-480-1160219  
; Sequence 1160219, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1160219
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1160219

Query Match      80.0%; Score 16; DB 12; Length 996;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 TTGGTTGGGCACACA 19
Db      823 TTGGTTGGGCACACA 838
|||||

RESULT 14
US-10-301-480-553142/c
; Sequence 553142, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 553142
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-553142

Query Match      80.0%; Score 16; DB 12; Length 996;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCTTTGGTTGGGCAAC 16
Db      40 GCTTTGGTTGGGCAAC 25
|||||

RESULT 15
US-10-301-480-553488/c
; Sequence 553488, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1160219
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1160219

Query Match      80.0%; Score 16; DB 12; Length 996;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCTTTGGTTGGGCAAC 16
Db      40 GCTTTGGTTGGGCAAC 25
|||||

RESULT 16
US-10-301-480-1166551/c
; Sequence 1166551, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1166551
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1166551

Query Match      80.0%; Score 16; DB 12; Length 996;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCTTTGGTTGGGCAAC 16
Db      40 GCTTTGGTTGGGCAAC 25
|||||

RESULT 17
US-10-301-480-1166897/c
; Sequence 1166897, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1166897
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1166897

Query Match      80.0%; Score 16; DB 12; Length 996;
```

```
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 553488
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-553488

Query Match      80.0%; Score 16; DB 12; Length 996;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCTTTGGTTGGGCAAC 16
Db      40 GCTTTGGTTGGGCAAC 25
|||||

RESULT 16
US-10-301-480-1166551/c
; Sequence 1166551, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1166551
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1166551

Query Match      80.0%; Score 16; DB 12; Length 996;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCTTTGGTTGGGCAAC 16
Db      40 GCTTTGGTTGGGCAAC 25
|||||

RESULT 17
US-10-301-480-1166897/c
; Sequence 1166897, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1166897
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1166897

Query Match      80.0%; Score 16; DB 12; Length 996;
```

```
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAAC 16
Db 40 GCTTTGGTTGGGCAAC 25

RESULT 18
US-11-136-527-34024/c
; Sequence 34024, Application US/11136527
; Publication No. US2005028750A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Mounts, William M
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34024
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Probe
US-11-136-527-34024

Query Match 75.0%; Score 15; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20
Db 25 GGTGGGCAACACAT 11

RESULT 19
US-10-301-480-541667
; Sequence 541667, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 541667
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-541667

Query Match 75.0%; Score 15; DB 12; Length 1000;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAAC 15
Db 691 GCTTTGGTTGGGCAAC 705

RESULT 20
US-10-301-480-541667
; Sequence 541667, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 541667
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-541667

Query Match 75.0%; Score 15; DB 12; Length 1000;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAAC 15
Db 691 GCTTTGGTTGGGCAAC 705

RESULT 21
US-09-925-065A-669805
; Sequence 669805, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669805
; LENGTH: 1153
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-669805

Query Match 75.0%; Score 15; DB 7; Length 1153;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTTGGTTGGGCAAC 16
Db 119 CTTTGGTTGGGCAAC 133

RESULT 22
US-10-115-609-22
; Sequence 22, Application US/10115609
; Publication No. US20060034826A1
```

```
US-10-301-480-1155076
; Sequence 1155076, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155076
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-1155076

Query Match 75.0%; Score 15; DB 12; Length 1000;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAAC 15
Db 691 GCTTTGGTTGGGCAAC 705

RESULT 21
US-09-925-065A-669805
; Sequence 669805, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669805
; LENGTH: 1153
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-669805

Query Match 75.0%; Score 15; DB 7; Length 1153;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTTGGTTGGGCAAC 16
Db 119 CTTTGGTTGGGCAAC 133

RESULT 22
US-10-115-609-22
; Sequence 22, Application US/10115609
; Publication No. US20060034826A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Carreno, Beatriz M.
; APPLICANT: Leonard, John
; TITLE OF INVENTION: Use of Agents That Modulate The Interaction Between Pd-1 And Its
; FILE REFERENCE: GNN-037
; CURRENT APPLICATION NUMBER: US/10/115,609
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 60/281,541
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 3593
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17)..(889)
US-10-115-609-22

Query Match 75.0%; Score 15; DB 11; Length 3593;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GGTGGGCAACACAT 20
DB 3406 GGTGGGCAACACAT 3420

RESULT 23
US-11-136-527-34011/c
; Sequence 34011, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34011
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Probe
US-11-136-527-34011

Query Match 70.0%; Score 14; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GTTGGGCAACACAT 20
DB 25 GTTGGGCAACACAT 12

RESULT 24
US-11-124-367A-25760/c
; Sequence 25760, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21

```

```

; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25760
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-367A-25760

Query Match 70.0%; Score 14; DB 17; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTGGTTGGCA 14
DB 67 GCTTGGTTGGCA 54

RESULT 25
US-09-925-065A-153801/c
; Sequence 153801, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153801
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-153801

Query Match 70.0%; Score 14; DB 7; Length 403;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTTTGGTTGGCAA 15
DB 259 CTTTGGTTGGCAA 246

RESULT 26
US-10-301-480-247396/c
; Sequence 247396, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21

```

; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 247396  
; LENGTH: 404  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-247396

Query Match 70.0%; Score 14; DB 12; Length 404;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTTGGTTGGCAA 15  
| | | | | | | | | |  
Db 260 CTTTGGTTGGCAA 247

## RESULT 27

US-10-301-480-860805/c  
; Sequence 860805, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 860805  
; LENGTH: 404  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-860805

Query Match 70.0%; Score 14; DB 12; Length 404;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTTGGTTGGCAA 15  
| | | | | | | | | |  
Db 260 CTTTGGTTGGCAA 247

## RESULT 28

US-10-301-480-14320/c  
; Sequence 14320, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14320  
; LENGTH: 468  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-603534

; ORGANISM: Homo sapien  
US-10-301-480-14320

Query Match 70.0%; Score 14; DB 11; Length 468;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TTGGTTGGCAACA 17  
| | | | | | | | | |  
Db 185 TTGGTTGGCAACA 172

## RESULT 29

US-10-301-480-627729/c  
; Sequence 627729, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 627729  
; LENGTH: 468  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-627729

Query Match 70.0%; Score 14; DB 12; Length 468;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TTGGTTGGCAACA 17  
| | | | | | | | | |  
Db 185 TTGGTTGGCAACA 172

## RESULT 30

US-09-925-065A-603534/c  
; Sequence 603534, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 603534  
; LENGTH: 506  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-603534

```
Query Match          70.0%; Score 14; DB 7; Length 506;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTTTGGTTGGCAA 15
DB 21 CTTTGGTTGGCAA 8

RESULT 31
US-09-925-065A-603533/c
; Sequence 603533, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 603533
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-603533

Query Match          70.0%; Score 14; DB 7; Length 515;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTTTGGTTGGCAA 15
DB 18 CTTTGGTTGGCAA 5

RESULT 32
US-10-301-480-425083
; Sequence 425083, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 425083
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-425083

Query Match          70.0%; Score 14; DB 12; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCTTTGGTTGGCA 14
DB 336 GCTTTGGTTGGCA 349

RESULT 33
US-10-301-480-425084
; Sequence 425084, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 425084
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-425084

Query Match          70.0%; Score 14; DB 12; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCA 14
DB 336 GCTTTGGTTGGCA 349

RESULT 34
US-10-301-480-425085
; Sequence 425085, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 425085
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-425085

Query Match          70.0%; Score 14; DB 12; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCA 14
DB 336 GCTTTGGTTGGCA 349

RESULT 35
US-10-301-480-1038492
; Sequence 1038492, Application US/10301480
```

```
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1038492
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1038492

Query Match      70.0%; Score 14; DB 12; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCTTTGGTTGGCA 14
      |||||||
Db      336 GCTTTGGTTGGCA 349

RESULT 36
US-10-301-480-1038493
; Sequence 1038493, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1038493
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1038493

Query Match      70.0%; Score 14; DB 12; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCTTTGGTTGGCA 14
      |||||||
Db      336 GCTTTGGTTGGCA 349

RESULT 37
US-10-301-480-1038494
; Sequence 1038494, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
```

```
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1038494
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1038494

Query Match      70.0%; Score 14; DB 12; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCTTTGGTTGGCA 14
      |||||||
Db      336 GCTTTGGTTGGCA 349

RESULT 38
US-09-925-065A-353424
; Sequence 353424, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 353424
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-353424

Query Match      70.0%; Score 14; DB 7; Length 537;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCTTTGGTTGGCA 14
      |||||||
Db      336 GCTTTGGTTGGCA 349

RESULT 39
US-09-925-065A-353425
; Sequence 353425, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
```



```
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 353425
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-353425
```

```
Query Match 70.0%; Score 14; DB 7; Length 537;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCTTTGGTTGGCA 14
|||||
DB 336 GCTTTGGTTGGCA 349
```

```
RESULT 40
US-09-925-065A-353426
; Sequence 353426, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 353426
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-353426
```

```
Query Match 70.0%; Score 14; DB 7; Length 537;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCTTTGGTTGGCA 14
|||||
DB 336 GCTTTGGTTGGCA 349
```

```
RESULT 41
US-09-925-065A-603532/c
; Sequence 603532, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
```

```
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 603532
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-603532
```

```
Query Match 70.0%; Score 14; DB 7; Length 538;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 CTTTGGTTGGCAA 15
|||||
DB 368 CTTTGGTTGGCAA 355
```

```
RESULT 42
US-09-925-065A-94527/c
; Sequence 94527, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94527
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-94527
```

```
Query Match 70.0%; Score 14; DB 7; Length 545;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 GTTTGGCAACACA 19
|||||
DB 109 GTTTGGCAACACA 96
```

```
RESULT 43
US-10-301-480-195769/c
; Sequence 195769, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
```

```
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195769
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-195769
```

```
Query Match 70.0%; Score 14; DB 11; Length 545;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 6 GGTGGGCAACACA 19
| | | | | | | | | |
Db 109 GGTGGGCAACACA 96
```

```
RESULT 44
US-10-301-480-809178/c
; Sequence 809178, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 809178
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-809178
```

```
Query Match 70.0%; Score 14; DB 12; Length 545;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 6 GGTGGGCAACACA 19
| | | | | | | | | |
Db 109 GGTGGGCAACACA 96
```

```
RESULT 45
US-10-301-480-415497/c
; Sequence 415497, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 415497
```

```
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-415497
```

```
Query Match 70.0%; Score 14; DB 12; Length 556;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 6 GGTGGGCAACACA 19
| | | | | | | | | |
Db 293 GGTGGGCAACACA 280
```

```
RESULT 46
US-10-301-480-415498/c
; Sequence 415498, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 415498
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-415498
```

```
Query Match 70.0%; Score 14; DB 12; Length 556;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 6 GGTGGGCAACACA 19
| | | | | | | | | |
Db 293 GGTGGGCAACACA 280
```

```
RESULT 47
US-10-301-480-1028906/c
; Sequence 1028906, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1028906
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1028906
```

```
Query Match 70.0%; Score 14; DB 12; Length 556;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 6 GGTGGGCAACACA 19  
Db 293 GGTGGGCAACACA 280

## RESULT 48

US-10-301-480-1028907/c  
; Sequence 1028907, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 10827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1028907  
; LENGTH: 556  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-1028907

Query Match 70.0%; Score 14; DB 12; Length 556;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GGTGGGCAACACA 19  
Db 293 GGTGGGCAACACA 280

## RESULT 49

US-09-925-065A-386479  
; Sequence 386479, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 10827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 386479  
; LENGTH: 562  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-386479

Query Match 70.0%; Score 14; DB 7; Length 562;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TTGGTTGGCAACA 17  
Db 58 TTGGTTGGCAACA 71

## RESULT 50

US-10-301-480-455055  
; Sequence 455055, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 10827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 455055  
; LENGTH: 564  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-455055

Query Match 70.0%; Score 14; DB 12; Length 564;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TTGGTTGGCAACA 17  
Db 58 TTGGTTGGCAACA 71

Search completed: May 8, 2006, 01:26:57  
Job time : 447 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2006, 22:42:33 ; Search time 1045 Seconds  
(without alignments)  
158.266 Million cell updates/sec

Title: US-10-671-074-172

Perfect score: 20

Sequence: 1 gcttggtgggaacacat 20

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 1

Total number of hits satisfying chosen parameters: 19584538

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA Main:\*

1: /cgm2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*

2: /cgm2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*

3: /cgm2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*

4: /cgm2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*

5: /cgm2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*

6: /cgm2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*

7: /cgm2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*

8: /cgm2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*

9: /cgm2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*

10: /cgm2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	US-10-671-074-172	Sequence 172, App
2	20	100.0	25	US-10-719-900-85415	Sequence 85415, A
3	20	100.0	254	US-09-794-928A-12	Sequence 12, Appl
4	20	100.0	303	US-10-242-535A-24143	Sequence 24143, A
5	20	100.0	303	US-10-085-783A-24143	Sequence 24143, A
6	20	100.0	447	US-10-242-535A-35649	Sequence 35649, A
7	20	100.0	447	US-10-085-783A-35649	Sequence 35649, A
8	20	100.0	520	US-10-505-680-708	Sequence 708, App
9	20	100.0	2514	US-10-450-763-20856	Sequence 20856, A
10	20	100.0	2827	US-10-391-530-1	Sequence 1, Appl
11	20	100.0	3200	US-10-391-530-3	Sequence 3, Appl
12	20	100.0	3517	US-10-391-530-2	Sequence 2, Appl
13	20	100.0	4945	US-10-671-074-11	Sequence 11, Appl
14	20	100.0	5723	US-10-007-926A-134	Sequence 134, App
15	20	100.0	5723	US-10-341-434-52	Sequence 52, Appl
16	20	100.0	5723	US-10-671-074-4	Sequence 4, Appl
17	20	100.0	5723	US-10-956-157-760	Sequence 760, App
18	20	100.0	5833	US-10-450-763-20857	Sequence 20857, A
19	19	95.0	20	US-10-671-074-40	Sequence 40, Appl
20	19	95.0	20	US-10-671-074-118	Sequence 118, App
21	17	85.0	2000	US-09-938-842A-4239	Sequence 4239, App
22	17	85.0	2000	US-09-938-842A-4239	Sequence 4239, App
23	16	80.0	20	US-10-671-074-41	Sequence 41, Appl

	20	7	US-10-671-074-119	Sequence 119, App
	4106	3	US-09-764-891-7739	Sequence 7739, App
	25	7	US-10-719-956-641503	Sequence 641503, A
	205	8	US-10-719-900-643422	Sequence 643422, A
	405	7	US-10-424-599-107214	Sequence 107214, A
	406	7	US-10-430-201-4459	Sequence 4459, App
	406	7	US-10-430-201-4460	Sequence 4460, App
	498	3	US-09-918-995-32485	Sequence 32485, A
	574	3	US-09-764-872-161	Sequence 161, App
	602	8	US-10-425-115-78789	Sequence 78789, A
	11153	4	US-09-925-065A-669805	Sequence 669805, A
	1263	9	US-10-450-763-11672	Sequence 11672, A
	1342	3	US-09-956-004-66	Sequence 66, Appl
	1342	6	US-10-808-570-66	Sequence 66, Appl
	1550	6	US-10-108-260A-1546	Sequence 1546, App
	212154	7	US-10-282-122A-9084	Sequence 9084, App
	3059	6	US-10-094-749-916	Sequence 916, App
	3342	10	US-11-097-143-32554	Sequence 32554, A
	3366	10	US-11-097-143-23078	Sequence 23078, A
	3593	5	US-10-002-775-10	Sequence 10, Appl
	3593	5	US-10-115-615-22	Sequence 22, Appl
	3593	9	US-10-764-420-1546	Sequence 1546, App
	3593	9	US-10-631-467-1140	Sequence 1140, App
	5205	8	US-10-758-672A-18	Sequence 18, Appl
	5205	9	US-10-758-636A-18	Sequence 18, Appl
	5466	8	US-10-357-819-9	Sequence 9, Appl
	5492	10	US-11-097-143-23077	Sequence 23077, A
	6308	8	US-10-758-672A-1	Sequence 1, Appl
	6308	9	US-10-758-636A-1	Sequence 1, Appl
	7742	6	US-10-287-218-22	Sequence 22, Appl
	7742	7	US-10-474-291-22	Sequence 22, Appl
	87766	9	US-10-795-159-685	Sequence 685, App
	130121	7	US-10-329-670-1	Sequence 1, Appl
	130121	8	US-10-158-865-1	Sequence 1, Appl
	130121	9	US-10-981-687-1	Sequence 1, Appl
	20	7	US-10-688-706-95	Sequence 95, Appl
	20	7	US-10-688-706-125	Sequence 125, App
	20	7	US-10-688-706-141	Sequence 141, App
	20	7	US-10-688-706-170	Sequence 170, App
	20	7	US-10-688-706-273	Sequence 273, App
	20	7	US-10-688-706-695	Sequence 695, App
	20	7	US-10-688-706-727	Sequence 727, App
	25	8	US-10-719-900-219203	Sequence 219203, A
	229	3	US-09-922-293-1203	Sequence 1203, App
	231	3	US-09-922-293-1201	Sequence 1201, App
	266	3	US-09-922-293-1220	Sequence 1220, App
	296	6	US-10-029-386-23544	Sequence 23544, A
	306	10	US-11-097-143-29267	Sequence 29267, A
	317	7	US-10-424-539-126752	Sequence 126752, A
	387	7	US-10-242-535A-33973	Sequence 33973, A
	387	7	US-10-085-783A-33973	Sequence 33973, A
	403	4	US-09-925-065A-153801	Sequence 153801, A
	404	3	US-09-922-293-1224	Sequence 1224, App
	419	3	US-09-922-293-1230	Sequence 1230, App
	426	3	US-09-783-590-10461	Sequence 10461, A
	440	3	US-09-866-050A-77	Sequence 77, Appl
	440	5	US-10-152-661-77	Sequence 77, Appl
	471	3	US-09-918-995-31180	Sequence 31180, A
	493	5	US-10-027-632-78715	Sequence 78715, A
	493	5	US-10-027-632-78715	Sequence 78715, A
	493	6	US-10-027-632-78715	Sequence 78715, A
	493	6	US-10-027-632-78715	Sequence 78715, A
	501	9	US-10-972-079-75011	Sequence 75011, A
	504	6	US-10-029-386-9844	Sequence 9844, App
	506	4	US-09-925-065A-603534	Sequence 603534, A
	515	4	US-09-925-065A-603533	Sequence 603533, A
	517	5	US-10-027-632-113357	Sequence 113357, A
	517	6	US-10-027-632-113357	Sequence 113357, A
	537	4	US-09-925-065A-353424	Sequence 353424, A
	537	4	US-09-925-065A-353425	Sequence 353425, A
	537	4	US-09-925-065A-353426	Sequence 353426, A
	538	4	US-09-925-065A-603532	Sequence 603532, A
	545	4	US-09-925-065A-94527	Sequence 94527, A

c 97	14	70.0	562	4	US-09-925-065A-386479	Sequence 386479,	c 170	14	70.0	2197	9	US-10-956-157-1519	Sequence 1519, Ap
c 98	14	70.0	567	4	US-09-925-065A-342764	Sequence 342764,	c 171	14	70.0	2197	9	US-10-756-149-169	Sequence 169, App
c 99	14	70.0	567	4	US-09-925-065A-342765	Sequence 342765,	c 172	14	70.0	2218	7	US-10-424-599-59900	Sequence 59900, A
100	14	70.0	587	4	US-09-925-065A-605689	Sequence 605689,	173	14	70.0	2227	8	US-10-723-860-7344	Sequence 7344, Ap
101	14	70.0	592	4	US-09-925-065A-199507	Sequence 199507,	c 174	14	70.0	2306	10	US-11-097-143-29266	Sequence 29266, A
102	14	70.0	592	4	US-09-925-065A-463575	Sequence 463575,	c 175	14	70.0	2314	5	US-10-027-632-261567	Sequence 261567,
103	14	70.0	592	4	US-09-925-065A-463576	Sequence 463576,	c 176	14	70.0	2314	5	US-10-027-632-261567	Sequence 261567,
104	14	70.0	592	4	US-09-925-065A-463577	Sequence 463577,	c 177	14	70.0	2314	6	US-10-027-632-261567	Sequence 261567,
105	14	70.0	592	4	US-09-925-065A-463578	Sequence 463578,	c 178	14	70.0	2314	6	US-10-027-632-261568	Sequence 261568,
106	14	70.0	593	7	US-10-021-323-14958	Sequence 14958, A	c 179	14	70.0	2393	7	US-10-236-115-577	Sequence 577, App
c 107	14	70.0	601	4	US-09-925-065A-540729	Sequence 540729,	c 180	14	70.0	2407	8	US-10-357-930-21308	Sequence 21308, A
c 108	14	70.0	602	4	US-09-925-065A-902618	Sequence 902618,	c 181	14	70.0	2407	8	US-10-357-930-23642	Sequence 23642, A
109	14	70.0	606	4	US-09-925-065A-199508	Sequence 199508,	c 182	14	70.0	2407	8	US-10-357-930-27150	Sequence 27150, A
110	14	70.0	611	4	US-09-925-065A-487106	Sequence 487106,	c 183	14	70.0	2407	8	US-10-357-930-29533	Sequence 29533, A
111	14	70.0	631	5	US-10-027-632-145308	Sequence 145308,	c 184	14	70.0	2407	8	US-10-450-763-23736	Sequence 23736, A
112	14	70.0	631	6	US-10-027-632-145308	Sequence 145308,	c 185	14	70.0	2439	7	US-10-052-482-228	Sequence 228, App
c 113	14	70.0	632	8	US-10-357-930-57240	Sequence 57240, A	c 186	14	70.0	2586	5	US-10-084-817-173	Sequence 173, App
c 114	14	70.0	646	5	US-10-027-632-68455	Sequence 68455, A	c 187	14	70.0	2793	9	US-10-486-970-1	Sequence 1, Appli
c 115	14	70.0	646	6	US-10-027-632-68455	Sequence 68455, A	c 188	14	70.0	2935	7	US-10-052-482-227	Sequence 227, App
c 116	14	70.0	661	4	US-09-925-065A-884132	Sequence 884132,	c 189	14	70.0	3036	6	US-10-369-493-26431	Sequence 26431, A
117	14	70.0	707	3	US-09-902-563-7	Sequence 7, Appli	c 190	14	70.0	3082	6	US-10-354-358-33	Sequence 33, Appl
118	14	70.0	707	5	US-10-096-255-7	Sequence 7, Appli	c 191	14	70.0	3082	7	US-10-688-706-3070	Sequence 3070, Ap
119	14	70.0	707	9	US-10-997-920-7	Sequence 7, Appli	c 192	14	70.0	3082	8	US-10-737-450-63	Sequence 63, Appl
120	14	70.0	723	9	US-10-504-328-5	Sequence 5, Appli	c 193	14	70.0	3082	8	US-10-753-267-11	Sequence 11, Appl
121	14	70.0	748	9	US-10-724-264A-30	Sequence 30, Appl	c 194	14	70.0	3082	8	US-10-831-704-68	Sequence 68, Appl
122	14	70.0	750	3	US-09-847-513A-8	Sequence 8, Appli	c 195	14	70.0	3082	9	US-10-956-157-767	Sequence 767, App
123	14	70.0	750	3	US-09-847-513A-14	Sequence 14, Appl	c 196	14	70.0	3082	9	US-10-956-157-767	Sequence 767, App
124	14	70.0	750	3	US-09-847-513A-26	Sequence 26, Appl	c 197	14	70.0	3136	7	US-10-756-149-1758	Sequence 1758, Ap
125	14	70.0	750	3	US-09-847-513A-28	Sequence 28, Appl	c 198	14	70.0	3236	4	US-10-262-511-121	Sequence 121, App
126	14	70.0	750	3	US-09-847-513A-36	Sequence 36, Appl	c 199	14	70.0	4169	10	US-09-925-065A-688254	Sequence 688254,
127	14	70.0	753	9	US-10-724-264A-10	Sequence 10, Appl	c 200	14	70.0	4963	9	US-11-097-143-40138	Sequence 40138, A
128	14	70.0	753	9	US-10-724-264A-19	Sequence 19, Appl	c 201	14	70.0	5030	10	US-11-097-143-19477	Sequence 19477, A
129	14	70.0	753	9	US-10-724-264A-20	Sequence 20, Appl	c 202	14	70.0	5358	10	US-11-097-143-13135	Sequence 13135, A
130	14	70.0	753	9	US-10-724-264A-22	Sequence 22, Appl	c 203	14	70.0	5403	3	US-09-902-563-3	Sequence 3, Appli
c 131	14	70.0	774	7	US-10-424-599-119148	Sequence 119148,	c 204	14	70.0	5403	5	US-10-096-255-3	Sequence 3, Appli
c 132	14	70.0	787	7	US-10-437-963-95022	Sequence 95022, A	c 205	14	70.0	5403	3	US-10-997-920-3	Sequence 3, Appli
c 133	14	70.0	816	6	US-10-369-493-27565	Sequence 27565, A	c 206	14	70.0	5502	3	US-09-349-058-45	Sequence 45, Appl
134	14	70.0	825	9	US-10-504-328-3	Sequence 3, Appli	c 207	14	70.0	5502	5	US-10-040-430-45	Sequence 45, Appl
135	14	70.0	887	3	US-09-764-877-2110	Sequence 2110, Ap	c 208	14	70.0	5555	5	US-10-450-763-29829	Sequence 29829, A
136	14	70.0	887	3	US-09-764-877-2112	Sequence 2112, Ap	c 209	14	70.0	8002	10	US-11-097-143-13348	Sequence 13348, A
137	14	70.0	887	6	US-10-242-515-2110	Sequence 2110, Ap	c 210	14	70.0	10271	3	US-09-754-468-43	Sequence 43, Appl
138	14	70.0	887	6	US-10-242-515-2112	Sequence 2112, Ap	c 211	14	70.0	14283	10	US-11-097-143-14938	Sequence 14938, A
139	14	70.0	921	3	US-09-764-877-2111	Sequence 2111, Ap	c 212	14	70.0	16995	2	US-08-961-527-82	Sequence 82, Appl
140	14	70.0	921	6	US-10-242-515-2111	Sequence 2111, Ap	c 213	14	70.0	16995	7	US-10-158-844-82	Sequence 82, Appl
141	14	70.0	1055	7	US-10-767-701-9453	Sequence 9453, Ap	c 214	14	70.0	24526	3	US-09-749-589-3	Sequence 3, Appli
142	14	70.0	1060	7	US-10-767-701-2116	Sequence 2116, Ap	c 215	14	70.0	24526	8	US-10-684-532-3	Sequence 3, Appli
143	14	70.0	1083	9	US-10-504-328-1	Sequence 1, Appli	c 216	14	70.0	29282	8	US-10-719-993-6786	Sequence 6786, Ap
c 144	14	70.0	1155	7	US-10-437-963-49492	Sequence 49492, A	c 217	14	70.0	50927	7	US-10-367-094-56	Sequence 56, Appl
c 145	14	70.0	1157	9	US-10-450-763-25810	Sequence 25810, A	c 218	14	70.0	52192	8	US-10-741-600-17615	Sequence 17615, A
146	14	70.0	1194	9	US-10-617-320-294	Sequence 294, App	c 219	14	70.0	89328	3	US-09-873-367C-332	Sequence 332, App
c 147	14	70.0	1279	3	US-09-765-272-47	Sequence 47, App	c 220	14	70.0	89328	9	US-10-843-641A-332	Sequence 332, App
c 148	14	70.0	1279	10	US-11-106-649-47	Sequence 47, Appl	c 221	14	70.0	96597	7	US-10-052-482-226	Sequence 226, App
c 149	14	70.0	1281	8	US-10-472-928-639	Sequence 639, App	c 222	14	70.0	104729	8	US-10-723-860-1434	Sequence 1434, Ap
c 150	14	70.0	1301	8	US-10-425-115-140395	Sequence 140395,	c 223	14	70.0	104729	9	US-10-756-149-1398	Sequence 1398, Ap
c 151	14	70.0	1308	9	US-10-617-320-1653	Sequence 1653, Ap	c 224	14	70.0	137908	9	US-10-287-436A-779	Sequence 779, App
c 152	14	70.0	1375	9	US-10-450-763-8896	Sequence 8896, Ap	c 225	14	70.0	201239	8	US-10-278-698-246	Sequence 246, App
c 153	14	70.0	1400	9	US-10-956-157-6754	Sequence 6754, Ap	c 226	14	70.0	201239	8	US-10-278-698-760	Sequence 760, App
c 154	14	70.0	1448	9	US-10-287-436A-952	Sequence 952, App	c 227	14	70.0	201239	8	US-10-472-928-4979	Sequence 4979, Ap
c 155	14	70.0	1449	8	US-10-723-860-3619	Sequence 3619, App	c 228	14	70.0	2162598	7	US-10-297-468A-1	Sequence 1, Appli
c 156	14	70.0	1492	8	US-10-767-795-1077	Sequence 1077, Ap	c 229	14	70.0	2731748	7	US-10-688-706-444	Sequence 444, App
c 157	14	70.0	1588	7	US-10-282-122A-12182	Sequence 12182, A	c 230	13	65.0	20	7	US-10-688-706-1058	Sequence 1058, Ap
c 158	14	70.0	1599	7	US-10-425-114-36073	Sequence 36073, A	c 231	13	65.0	20	7	US-10-719-956-120333	Sequence 120333,
c 159	14	70.0	1614	9	US-10-450-763-26109	Sequence 26109, A	c 232	13	65.0	25	7	US-10-719-956-454716	Sequence 454716,
c 160	14	70.0	1690	7	US-10-767-701-14451	Sequence 14451, A	c 233	13	65.0	25	8	US-10-719-900-655350	Sequence 655350,
c 161	14	70.0	1931	4	US-09-925-065A-723533	Sequence 723533,	c 234	13	65.0	25	8	US-10-719-900-687374	Sequence 687374,
c 162	14	70.0	1992	9	US-10-450-763-5168	Sequence 5168, Ap	c 235	13	65.0	25	8	US-10-719-900-851032	Sequence 851032,
c 163	14	70.0	2029	7	US-10-297-599-1	Sequence 1, Appli	c 236	13	65.0	25	8	US-10-719-900-907847	Sequence 907847,
c 164	14	70.0	2129	9	US-10-486-970-9	Sequence 9, Appli	c 237	13	65.0	25	10	US-11-036-317-121416	Sequence 121416,
c 165	14	70.0	2179	7	US-10-504-173-112	Sequence 112, App	c 238	13	65.0	25	10	US-11-036-317-495161	Sequence 495161,
c 166	14	70.0	2179	9	US-10-437-963-49495	Sequence 49495, A	c 239	13	65.0	25	10	US-11-036-317-555889	Sequence 555889,
c 167	14	70.0	2196	4	US-09-925-065A-683172	Sequence 683172,	c 240	13	65.0	25	10	US-11-036-317-721238	Sequence 721238,
c 168	14	70.0	2197	6	US-10-172-118-1228	Sequence 1228, Ap	c 241	13	65.0	60	3	US-09-808-975-11722	Sequence 11722, A
c 169	14	70.0	2197	7	US-10-342-887-1228	Sequence 1228, Ap	c 242	13	65.0	143	3	US-09-815-242-2038	Sequence 2038, Ap

C 243	13	65.0	143	7	US-10-282-122A-4568	Sequence 4568, Ap	C 316	13	65.0	566	7	US-10-767-701-1060	Sequence 1060, Ap
C 244	13	65.0	161	3	US-09-983-646A-4	Sequence 4, Appli	C 317	13	65.0	567	3	US-09-796-632-4178	Sequence 4178, Ap
C 245	13	65.0	168	3	US-09-747-377-54	Sequence 54, Appli	C 318	13	65.0	567	5	US-10-040-862-4178	Sequence 4178, Ap
C 246	13	65.0	201	7	US-10-105-613-54	Sequence 54, Appli	C 319	13	65.0	567	6	US-10-057-478B-4178	Sequence 4178, Ap
C 247	13	65.0	201	7	US-10-741-601-25881	Sequence 25881, A	C 320	13	65.0	567	6	US-10-154-894B-4178	Sequence 4178, Ap
C 248	13	65.0	201	7	US-10-741-601-25909	Sequence 25909, A	C 321	13	65.0	567	8	US-10-764-324-4178	Sequence 4178, Ap
C 249	13	65.0	201	8	US-10-719-993-34294	Sequence 34294, A	C 322	13	65.0	572	5	US-10-027-632-209217	Sequence 209217, A
C 250	13	65.0	216	6	US-10-339-740-208	Sequence 208, App	C 323	13	65.0	572	5	US-10-027-632-218643	Sequence 218643, A
C 251	13	65.0	231	7	US-10-437-963-21400	Sequence 21400, A	C 324	13	65.0	572	6	US-10-027-632-209217	Sequence 209217, A
C 252	13	65.0	258	7	US-10-424-599-1814	Sequence 1814, Ap	C 325	13	65.0	572	6	US-10-027-632-218643	Sequence 218643, A
C 253	13	65.0	258	7	US-10-767-701-11706	Sequence 17706, A	C 326	13	65.0	572	9	US-10-972-079-68767	Sequence 68767, A
C 254	13	65.0	277	7	US-10-424-599-14939	Sequence 14939, A	C 327	13	65.0	574	7	US-10-767-701-23865	Sequence 23865, A
C 255	13	65.0	281	7	US-10-424-599-715	Sequence 715, App	C 328	13	65.0	577	8	US-10-425-115-42360	Sequence 42360, A
C 256	13	65.0	289	3	US-09-867-701-7257	Sequence 7257, Ap	C 329	13	65.0	579	7	US-10-437-963-6891	Sequence 6891, Ap
C 257	13	65.0	300	9	US-10-779-543-6238	Sequence 6238, Ap	C 330	13	65.0	580	6	US-10-029-386-1698	Sequence 1698, Ap
C 258	13	65.0	322	7	US-10-424-599-9127	Sequence 9127, Ap	C 331	13	65.0	588	4	US-09-925-065A-521903	Sequence 521903, A
C 259	13	65.0	340	8	US-10-425-115-111286	Sequence 11286, A	C 332	13	65.0	590	4	US-09-925-065A-246451	Sequence 246451, A
C 260	13	65.0	350	7	US-10-424-599-5044	Sequence 5044, Ap	C 333	13	65.0	591	4	US-09-925-065A-648059	Sequence 648059, A
C 261	13	65.0	353	3	US-09-815-249-21139	Sequence 2139, Ap	C 334	13	65.0	591	4	US-09-925-065A-935924	Sequence 935924, A
C 262	13	65.0	353	7	US-10-282-122A-4704	Sequence 4704, Ap	C 335	13	65.0	591	4	US-09-925-065A-935925	Sequence 935925, A
C 263	13	65.0	377	3	US-09-864-408A-5619	Sequence 5619, Ap	C 336	13	65.0	592	4	US-09-925-065A-335272	Sequence 335272, A
C 264	13	65.0	377	4	US-09-925-065A-491370	Sequence 491370, A	C 337	13	65.0	595	9	US-10-972-079-86171	Sequence 86171, A
C 265	13	65.0	387	8	US-10-425-115-866	Sequence 866, App	C 338	13	65.0	596	4	US-09-925-065A-754854	Sequence 754854, A
C 266	13	65.0	394	8	US-10-425-115-103841	Sequence 103841, A	C 339	13	65.0	596	4	US-09-925-065A-754855	Sequence 754855, A
C 267	13	65.0	398	7	US-10-437-963-56425	Sequence 56425, A	C 340	13	65.0	596	4	US-09-925-065A-825513	Sequence 825513, A
C 268	13	65.0	400	2	US-08-781-986A-2971	Sequence 2971, Ap	C 341	13	65.0	598	5	US-10-027-632-248084	Sequence 248084, A
C 269	13	65.0	400	7	US-10-329-624-2971	Sequence 2971, Ap	C 342	13	65.0	598	5	US-10-027-632-248085	Sequence 248085, A
C 270	13	65.0	401	8	US-10-425-115-108634	Sequence 108634, A	C 343	13	65.0	598	6	US-10-027-632-248084	Sequence 248084, A
C 271	13	65.0	404	7	US-10-424-599-49377	Sequence 49377, A	C 344	13	65.0	598	6	US-10-027-632-248085	Sequence 248085, A
C 272	13	65.0	407	8	US-10-425-115-56471	Sequence 56471, A	C 345	13	65.0	599	5	US-10-027-632-45269	Sequence 45269, A
C 273	13	65.0	410	7	US-10-424-599-15645	Sequence 15645, A	C 346	13	65.0	599	5	US-10-027-632-92411	Sequence 92411, A
C 274	13	65.0	416	7	US-10-424-599-100021	Sequence 100021, A	C 347	13	65.0	599	5	US-10-027-632-307601	Sequence 307601, A
C 275	13	65.0	418	3	US-09-560-863-555	Sequence 555, App	C 348	13	65.0	599	6	US-10-027-632-45269	Sequence 45269, A
C 276	13	65.0	423	7	US-10-437-963-85777	Sequence 85777, A	C 349	13	65.0	599	6	US-10-027-632-92411	Sequence 92411, A
C 277	13	65.0	425	4	US-09-925-065A-459665	Sequence 459665, A	C 350	13	65.0	599	6	US-10-027-632-307601	Sequence 307601, A
C 278	13	65.0	438	8	US-10-425-115-183222	Sequence 183222, A	C 351	13	65.0	600	9	US-10-972-079-26509	Sequence 26509, A
C 279	13	65.0	445	6	US-10-914-037-17	Sequence 17, Appli	C 352	13	65.0	600	9	US-10-972-079-26510	Sequence 26510, A
C 280	13	65.0	456	6	US-10-369-493-43615	Sequence 43615, A	C 353	13	65.0	600	9	US-10-972-079-26511	Sequence 26511, A
C 281	13	65.0	486	8	US-10-914-037-18	Sequence 18, Appli	C 354	13	65.0	600	9	US-10-972-079-26512	Sequence 26512, A
C 282	13	65.0	487	5	US-10-027-632-11192	Sequence 11192, A	C 355	13	65.0	600	9	US-10-972-079-28516	Sequence 28516, A
C 283	13	65.0	487	6	US-10-027-632-11192	Sequence 11192, A	C 356	13	65.0	600	9	US-10-972-079-43227	Sequence 43227, A
C 284	13	65.0	488	8	US-10-425-115-126141	Sequence 126141, A	C 357	13	65.0	600	9	US-10-972-079-43228	Sequence 43228, A
C 285	13	65.0	495	5	US-10-060-036-4296	Sequence 4296, Ap	C 358	13	65.0	600	9	US-10-972-079-44504	Sequence 44504, A
C 286	13	65.0	495	8	US-10-425-115-22728	Sequence 22728, A	C 359	13	65.0	600	9	US-10-972-079-88827	Sequence 88827, A
C 287	13	65.0	499	7	US-10-424-599-11028	Sequence 11028, A	C 360	13	65.0	600	9	US-09-925-065A-96029	Sequence 96029, A
C 288	13	65.0	511	4	US-09-925-065A-60422	Sequence 60422, A	C 361	13	65.0	609	4	US-09-925-065A-478182	Sequence 478182, A
C 289	13	65.0	515	4	US-09-925-065A-590910	Sequence 590910, A	C 362	13	65.0	609	4	US-09-925-065A-478183	Sequence 478183, A
C 290	13	65.0	520	4	US-09-925-065A-545704	Sequence 545704, A	C 363	13	65.0	613	5	US-10-027-632-130886	Sequence 130886, A
C 291	13	65.0	520	4	US-09-925-065A-545705	Sequence 545705, A	C 364	13	65.0	613	5	US-10-027-632-130887	Sequence 130887, A
C 292	13	65.0	521	5	US-10-027-632-89361	Sequence 89361, A	C 365	13	65.0	613	5	US-10-027-632-130888	Sequence 130888, A
C 293	13	65.0	521	5	US-10-027-632-89362	Sequence 89362, A	C 366	13	65.0	613	6	US-10-027-632-130886	Sequence 130886, A
C 294	13	65.0	521	6	US-10-027-632-89361	Sequence 89361, A	C 367	13	65.0	613	6	US-10-027-632-130887	Sequence 130887, A
C 295	13	65.0	521	6	US-10-027-632-89362	Sequence 89362, A	C 368	13	65.0	613	6	US-10-027-632-130888	Sequence 130888, A
C 296	13	65.0	526	6	US-10-029-386-844	Sequence 844, App	C 369	13	65.0	616	4	US-09-925-065A-471520	Sequence 471520, A
C 297	13	65.0	528	4	US-09-925-065A-224786	Sequence 224786, A	C 370	13	65.0	617	6	US-10-204-836-3	Sequence 3, Appli
C 298	13	65.0	529	3	US-09-920-300A-1580	Sequence 1580, Ap	C 371	13	65.0	621	4	US-09-925-065A-481726	Sequence 481726, A
C 299	13	65.0	529	5	US-10-033-528-1580	Sequence 1580, Ap	C 372	13	65.0	621	4	US-09-925-065A-481727	Sequence 481727, A
C 300	13	65.0	529	6	US-10-099-936-1580	Sequence 1580, Ap	C 373	13	65.0	623	7	US-10-437-963-6019	Sequence 6019, Ap
C 301	13	65.0	529	9	US-10-961-537-1580	Sequence 1580, Ap	C 374	13	65.0	625	5	US-10-027-632-82523	Sequence 82523, A
C 302	13	65.0	540	7	US-10-437-963-90426	Sequence 90426, A	C 375	13	65.0	625	5	US-10-027-632-302200	Sequence 302200, A
C 303	13	65.0	542	4	US-09-925-065A-755300	Sequence 755300, A	C 376	13	65.0	625	6	US-10-027-632-82523	Sequence 82523, A
C 304	13	65.0	545	4	US-09-925-065A-471519	Sequence 471519, A	C 377	13	65.0	625	6	US-10-027-632-302200	Sequence 302200, A
C 305	13	65.0	545	4	US-09-925-065A-471521	Sequence 471521, A	C 378	13	65.0	631	4	US-09-925-065A-411439	Sequence 411439, A
C 306	13	65.0	545	4	US-09-925-065A-471522	Sequence 471522, A	C 379	13	65.0	633	5	US-10-027-632-207358	Sequence 207358, A
C 307	13	65.0	546	4	US-09-925-065A-95641	Sequence 95641, A	C 380	13	65.0	633	6	US-10-027-632-207358	Sequence 207358, A
C 308	13	65.0	546	4	US-09-925-065A-95642	Sequence 95642, A	C 381	13	65.0	634	4	US-09-925-065A-49102	Sequence 49102, A
C 309	13	65.0	547	4	US-09-925-065A-507101	Sequence 507101, A	C 382	13	65.0	647	4	US-09-925-065A-512632	Sequence 512632, A
C 310	13	65.0	555	4	US-09-925-065A-333718	Sequence 333718, A	C 383	13	65.0	647	4	US-09-925-065A-512633	Sequence 512633, A
C 311	13	65.0	558	4	US-09-925-065A-640608	Sequence 640608, A	C 384	13	65.0	647	5	US-10-027-632-102334	Sequence 102334, A
C 312	13	65.0	558	7	US-10-282-132A-22137	Sequence 22137, A	C 385	13	65.0	647	6	US-10-027-632-102334	Sequence 102334, A
C 313	13	65.0	558	7	US-10-424-599-11211	Sequence 11211, A	C 386	13	65.0	655	3	US-09-864-408A-4283	Sequence 4283, Ap
C 314	13	65.0	559	4	US-09-925-065A-191767	Sequence 191767, A	C 387	13	65.0	655	4	US-09-925-065A-763741	Sequence 763741, A
C 315	13	65.0	559	5	US-10-060-036-1381	Sequence 1381, Ap	C 388	13	65.0	655	4	US-09-925-065A-763742	Sequence 763742, A

C 389	13	65.0	656	4	US-09-925-065A-771621, Sequence 771621,	C 462	13	65.0	1161	2	US-08-781-986A-497	Sequence 497, App
C 390	13	65.0	656	4	US-09-925-065A-771622, Sequence 771622,	C 463	13	65.0	1161	7	US-10-329-624-447	Sequence 497, App
C 391	13	65.0	658	4	US-09-925-065A-882868 Sequence 882868,	C 464	13	65.0	1163	7	US-10-425-114-22074	Sequence 20074, A
C 392	13	65.0	658	4	US-09-925-065A-889432 Sequence 889432,	C 465	13	65.0	1186	8	US-10-739-930-4848	Sequence 4848, Ap
C 393	13	65.0	666	8	US-10-425-115-10304, Sequence 10304, A	C 466	13	65.0	1193	7	US-10-437-963-10611	Sequence 10611, A
C 394	13	65.0	670	7	US-10-767-701-26141, Sequence 26141, A	C 467	13	65.0	1200	4	US-09-925-065A-2692	Sequence 2692, Ap
C 395	13	65.0	683	4	US-09-925-065A-79305, Sequence 79305, A	C 468	13	65.0	1200	4	US-09-925-065A-2693	Sequence 2693, Ap
C 396	13	65.0	692	5	US-10-027-632-99365 Sequence 99365, A	C 469	13	65.0	1226	7	US-10-398-221-1123	Sequence 1123, Ap
C 397	13	65.0	692	5	US-10-027-632-99366 Sequence 99366, A	C 470	13	65.0	1245	5	US-10-027-632-207357	Sequence 207357,
C 398	13	65.0	692	6	US-10-027-632-99366 Sequence 99366, A	C 471	13	65.0	1245	5	US-10-027-632-207357	Sequence 207357,
C 399	13	65.0	692	6	US-10-027-632-99366 Sequence 99366, A	C 472	13	65.0	1254	7	US-10-437-963-19139	Sequence 19139, A
C 400	13	65.0	694	6	US-10-149-759-67 Sequence 67, Appl	C 473	13	65.0	1277	7	US-10-425-114-32433	Sequence 32433, A
C 401	13	65.0	713	4	US-09-925-065A-82461 Sequence 82461, A	C 474	13	65.0	1281	9	US-10-617-320-1849	Sequence 1849, Ap
C 402	13	65.0	714	7	US-10-767-701-9488 Sequence 9488, Ap	C 475	13	65.0	1285	6	US-10-425-115-117753	Sequence 117753,
C 403	13	65.0	717	10	US-11-097-143-24035 Sequence 24035, A	C 476	13	65.0	1288	6	US-10-017-161-2017	Sequence 2017, Ap
C 404	13	65.0	725	4	US-09-925-065A-80813 Sequence 80813, A	C 477	13	65.0	1288	6	US-10-292-798-1663	Sequence 1663, Ap
C 405	13	65.0	725	4	US-09-925-065A-80813 Sequence 80813, A	C 478	13	65.0	1310	8	US-10-425-115-160183	Sequence 160183,
C 406	13	65.0	736	4	US-09-925-065A-932077 Sequence 932077, A	C 479	13	65.0	1316	9	US-10-450-763-5185	Sequence 5185, Ap
C 407	13	65.0	741	6	US-10-369-493-33020 Sequence 33020, A	C 480	13	65.0	1343	6	US-10-012-697-1340	Sequence 1340, Ap
C 408	13	65.0	741	6	US-10-369-493-33172 Sequence 33172, A	C 481	13	65.0	1343	9	US-10-779-543-23340	Sequence 23340, A
C 409	13	65.0	750	6	US-10-369-493-41196 Sequence 41196, A	C 482	13	65.0	1348	6	US-10-120-988-404	Sequence 404, App
C 410	13	65.0	752	4	US-09-925-065A-84226 Sequence 84226, A	C 483	13	65.0	1380	6	US-10-369-493-44726	Sequence 44726, A
C 411	13	65.0	752	6	US-10-012-697-651 Sequence 651, App	C 484	13	65.0	1386	7	US-10-282-122A-34123	Sequence 34123, A
C 412	13	65.0	752	9	US-10-779-543-22651 Sequence 22651, A	C 485	13	65.0	1418	4	US-09-925-065A-58680	Sequence 58680, A
C 413	13	65.0	771	7	US-10-767-701-14188 Sequence 14188, A	C 486	13	65.0	1454	6	US-10-369-493-25184	Sequence 25184, A
C 414	13	65.0	772	5	US-10-027-632-29156 Sequence 29156, A	C 487	13	65.0	1475	7	US-10-371-472-30	Sequence 30, Appl
C 415	13	65.0	772	6	US-10-027-632-29156 Sequence 29156, A	C 488	13	65.0	1475	7	US-10-371-472-30	Sequence 30, Appl
C 416	13	65.0	788	4	US-09-925-065A-69475 Sequence 69475, A	C 489	13	65.0	1491	9	US-10-450-763-9000	Sequence 9000, Ap
C 417	13	65.0	788	4	US-09-925-065A-69475 Sequence 69475, A	C 490	13	65.0	1502	7	US-10-072-012-23	Sequence 23, Appl
C 418	13	65.0	806	5	US-10-027-632-29717 Sequence 29717, A	C 491	13	65.0	1548	8	US-10-739-930-1177	Sequence 1177, Ap
C 419	13	65.0	806	6	US-10-027-632-29717 Sequence 29717, A	C 492	13	65.0	1584	6	US-10-369-493-25422	Sequence 25422, A
C 420	13	65.0	807	8	US-10-425-115-149036 Sequence 149036, A	C 493	13	65.0	1596	7	US-10-260-238-3314	Sequence 3314, Ap
C 421	13	65.0	810	4	US-09-925-065A-86355 Sequence 86355, A	C 494	13	65.0	1625	8	US-10-866-527-17	Sequence 17, Appl
C 422	13	65.0	830	7	US-10-425-114-5036 Sequence 5036, Ap	C 495	13	65.0	1626	8	US-10-866-527-21	Sequence 21, Appl
C 423	13	65.0	856	9	US-10-439-353A-96 Sequence 96, Appl	C 496	13	65.0	1641	7	US-10-425-114-28591	Sequence 28591, A
C 424	13	65.0	857	6	US-10-012-697-715, Sequence 715, App	C 497	13	65.0	1657	8	US-10-425-115-170674	Sequence 170674,
C 425	13	65.0	857	9	US-10-779-543-22715 Sequence 22715, A	C 498	13	65.0	1744	7	US-10-424-599-121358	Sequence 121358,
C 426	13	65.0	876	4	US-09-925-065A-715791 Sequence 715791, A	C 499	13	65.0	1803	3	US-09-815-242-4583	Sequence 4583, Ap
C 427	13	65.0	876	4	US-09-925-065A-715792 Sequence 715792, A	C 500	13	65.0	1806	7	US-10-282-122A-7863	Sequence 7863, Ap
C 428	13	65.0	876	4	US-09-925-065A-715793 Sequence 715793, A	C 501	13	65.0	1806	7	US-10-282-122A-31664	Sequence 31664, A
C 429	13	65.0	890	7	US-10-425-114-31415 Sequence 31415, A	C 502	13	65.0	1806	8	US-10-857-625-159	Sequence 159, App
C 430	13	65.0	903	8	US-10-425-115-13115 Sequence 13115, A	C 503	13	65.0	1806	8	US-10-857-625-152	Sequence 152, App
C 431	13	65.0	904	3	US-09-452-599-162 Sequence 162, App	C 504	13	65.0	1818	6	US-10-259-194A-181	Sequence 181, App
C 432	13	65.0	918	6	US-10-121-120-162 Sequence 162, App	C 505	13	65.0	1824	3	US-09-815-242-8270	Sequence 8270, Ap
C 433	13	65.0	918	6	US-10-121-120-162 Sequence 162, App	C 506	13	65.0	1824	3	US-09-815-242-8674	Sequence 8674, Ap
C 434	13	65.0	918	6	US-10-121-120-162 Sequence 162, App	C 507	13	65.0	1836	7	US-10-282-122A-8759	Sequence 8759, Ap
C 435	13	65.0	945	8	US-10-425-115-110999 Sequence 110999, A	C 508	13	65.0	1888	4	US-09-925-065A-707588	Sequence 707588,
C 436	13	65.0	945	8	US-10-774-355A-753 Sequence 753, App	C 509	13	65.0	1902	7	US-10-437-963-64766	Sequence 64766, A
C 437	13	65.0	963	7	US-10-424-599-47042 Sequence 47042, A	C 510	13	65.0	1927	7	US-10-398-221-2598	Sequence 2598, A
C 438	13	65.0	994	6	US-10-017-161-575 Sequence 138, App	C 511	13	65.0	1995	3	US-09-771-161A-31	Sequence 31, Appl
C 439	13	65.0	1007	8	US-10-739-930-138 Sequence 351, App	C 512	13	65.0	2000	3	US-10-425-115-10856	Sequence 10856, A
C 440	13	65.0	1014	7	US-10-389-647-351 Sequence 126576, A	C 513	13	65.0	2000	3	US-09-938-842A-4728	Sequence 4728, Ap
C 441	13	65.0	1015	8	US-10-425-115-126576 Sequence 126576, A	C 514	13	65.0	2000	3	US-09-938-842A-4728	Sequence 4728, Ap
C 442	13	65.0	1036	7	US-10-425-114-34263 Sequence 34263, A	C 515	13	65.0	2000	3	US-10-424-599-38425	Sequence 38425, A
C 443	13	65.0	1039	5	US-10-027-632-323709 Sequence 323709, A	C 516	13	65.0	2001	10	US-11-097-143-23570	Sequence 23570, A
C 444	13	65.0	1039	5	US-10-027-632-323726 Sequence 323726, A	C 517	13	65.0	2009	4	US-09-925-065A-82092	Sequence 82092, A
C 445	13	65.0	1039	5	US-10-027-632-323832 Sequence 323832, A	C 518	13	65.0	2009	4	US-09-925-065A-82093	Sequence 82093, A
C 446	13	65.0	1039	5	US-10-027-632-323909 Sequence 323909, A	C 519	13	65.0	2009	4	US-09-925-065A-82094	Sequence 82094, A
C 447	13	65.0	1039	6	US-10-027-632-323709 Sequence 323709, A	C 520	13	65.0	2031	7	US-10-398-221-664	Sequence 664, App
C 448	13	65.0	1039	6	US-10-027-632-323726 Sequence 323726, A	C 521	13	65.0	2083	4	US-09-925-065A-547048	Sequence 547048,
C 449	13	65.0	1039	6	US-10-027-632-323832 Sequence 323832, A	C 522	13	65.0	2083	5	US-10-027-632-258567	Sequence 258567,
C 450	13	65.0	1039	6	US-10-027-632-323909 Sequence 323909, A	C 523	13	65.0	2083	6	US-10-027-632-258567	Sequence 258567,
C 451	13	65.0	1041	7	US-10-437-963-80259 Sequence 80259, A	C 524	13	65.0	2089	3	US-09-764-864-1598	Sequence 1598, Ap
C 452	13	65.0	1063	6	US-10-085-198-179 Sequence 179, App	C 525	13	65.0	2094	8	US-10-425-115-149035	Sequence 149035,
C 453	13	65.0	1092	3	US-09-778-844-90 Sequence 90, Appl	C 526	13	65.0	2099	8	US-10-425-115-87145	Sequence 87145, A
C 454	13	65.0	1092	5	US-10-021-657-5 Sequence 5, Appl1	C 527	13	65.0	2100	8	US-10-425-115-10745	Sequence 10745, A
C 455	13	65.0	1092	6	US-10-412-000-5 Sequence 5, Appl1	C 528	13	65.0	2178	8	US-10-739-930-2282	Sequence 2282, Ap
C 456	13	65.0	1118	4	US-09-925-065A-71947 Sequence 71947, A	C 529	13	65.0	2311	3	US-09-800-729-66	Sequence 66, Appl
C 457	13	65.0	1118	4	US-09-925-065A-71948 Sequence 71948, A	C 530	13	65.0	2311	9	US-10-510-871-95	Sequence 95, Appl
C 458	13	65.0	1129	8	US-10-343-903-31 Sequence 31, Appl	C 531	13	65.0	2354	9	US-10-450-763-4409	Sequence 4409, Ap
C 459	13	65.0	1140	7	US-10-282-122A-34179 Sequence 34179, A	C 532	13	65.0	2357	10	US-11-097-143-35107	Sequence 35107, A
C 460	13	65.0	1155	3	US-09-815-242-8546 Sequence 8546, Ap	C 533	13	65.0	2388	3	US-09-933-767-154	Sequence 154, App
C 461	13	65.0	1155	3	US-09-815-242-8785 Sequence 8785, Ap	C 534	13	65.0	2388	5	US-10-004-860-154	Sequence 154, App



535	13	65.0	2388	5	US-10-023-282-154	Sequence 154, App	c 608	13	65.0	7239	5	US-10-084-817-195	Sequence 195, App
c 536	13	65.0	2394	3	US-09-800-729-33	Sequence 33, Appl	c 609	13	65.0	7555	10	US-11-097-143-12403	Sequence 12403, A
c 537	13	65.0	2465	10	US-11-097-143-6122	Sequence 6122, Ap	c 610	13	65.0	7911	10	US-11-097-143-6790	Sequence 6790, Ap
538	13	65.0	2538	8	US-10-425-115-67067	Sequence 67067, A	c 611	13	65.0	8094	10	US-11-097-143-2992	Sequence 2992, Ap
539	13	65.0	2598	9	US-10-450-763-21262	Sequence 21262, A	c 612	13	65.0	8955	8	US-10-857-625-51	Sequence 51, Appl
c 540	13	65.0	2631	3	US-09-801-368-43	Sequence 43, Appl	c 613	13	65.0	10200	10	US-11-097-143-19174	Sequence 19174, A
c 541	13	65.0	2661	5	US-10-027-633-263854	Sequence 263854, A	c 614	13	65.0	10645	10	US-11-097-143-16688	Sequence 16688, A
c 542	13	65.0	2661	5	US-10-027-633-263855	Sequence 263855, A	c 615	13	65.0	11451	8	US-10-857-625-95	Sequence 95, Appl
c 543	13	65.0	2661	6	US-10-027-633-263854	Sequence 263854, A	c 616	13	65.0	12138	10	US-11-097-143-20566	Sequence 20566, A
c 544	13	65.0	2661	6	US-10-027-633-263855	Sequence 263855, A	c 617	13	65.0	12161	10	US-11-097-143-20554	Sequence 20554, A
c 545	13	65.0	2717	10	US-11-097-143-24034	Sequence 24034, A	c 618	13	65.0	12332	10	US-11-097-143-13561	Sequence 13561, A
546	13	65.0	2850	9	US-10-481-032A-559	Sequence 559, App	c 619	13	65.0	14175	3	US-09-764-868-1474	Sequence 1474, Ap
547	13	65.0	2905	4	US-09-925-065A-712237	Sequence 712237, App	c 620	13	65.0	14396	10	US-11-097-143-9031	Sequence 9031, Ap
548	13	65.0	2905	4	US-09-925-065A-712238	Sequence 712238, App	c 621	13	65.0	20099	5	US-10-177-744A-12	Sequence 12, Appl
549	13	65.0	2905	4	US-09-925-065A-712239	Sequence 712239, App	c 622	13	65.0	21407	10	US-11-097-143-40579	Sequence 40579, A
550	13	65.0	2905	4	US-09-925-065A-712240	Sequence 712240, A	c 623	13	65.0	21732	3	US-09-764-872-717	Sequence 717, App
c 551	13	65.0	2930	7	US-10-437-963-26536	Sequence 26536, A	c 624	13	65.0	30306	5	US-10-087-192-523	Sequence 523, App
552	13	65.0	3024	4	US-09-925-065A-711172	Sequence 711172, A	c 625	13	65.0	32503	10	US-11-097-143-16687	Sequence 16687, A
553	13	65.0	3024	4	US-09-925-065A-711173	Sequence 711173, A	c 626	13	65.0	35042	9	US-10-915-740A-2	Sequence 2, Appli
554	13	65.0	3309	10	US-11-097-143-6791	Sequence 6791, Ap	c 627	13	65.0	35962	8	US-10-775-169-234	Sequence 234, App
c 555	13	65.0	3366	3	US-09-925-301-246	Sequence 246, App	c 628	13	65.0	35962	8	US-10-473-126-2	Sequence 2, Appli
c 556	13	65.0	3430	7	US-10-437-963-101945	Sequence 101945, A	c 629	13	65.0	44325	3	US-09-997-722-226	Sequence 226, App
557	13	65.0	3477	7	US-10-437-963-75383	Sequence 75383, A	c 630	13	65.0	45588	9	US-10-893-315-180	Sequence 180, App
558	13	65.0	3601	4	US-09-925-065A-81422	Sequence 81422, A	c 631	13	65.0	51193	9	US-10-737-082-46	Sequence 46, Appl
559	13	65.0	3601	4	US-09-925-065A-81423	Sequence 81423, A	c 632	13	65.0	51193	9	US-10-737-082-46	Sequence 46, Appl
560	13	65.0	3601	4	US-09-925-065A-81424	Sequence 81424, A	c 633	13	65.0	54039	8	US-10-741-600-17580	Sequence 17580, A
561	13	65.0	3601	4	US-09-925-065A-81425	Sequence 81425, A	c 634	13	65.0	55942	7	US-10-741-601-5775	Sequence 5775, Ap
562	13	65.0	3665	3	US-09-925-300-756	Sequence 756, App	c 635	13	65.0	61197	5	US-10-087-192-1924	Sequence 1924, Ap
c 563	13	65.0	3697	10	US-11-097-143-10675	Sequence 10675, A	c 636	13	65.0	61588	5	US-10-087-192-1570	Sequence 1570, Ap
564	13	65.0	3897	5	US-10-021-657-7	Sequence 7, Appli	c 637	13	65.0	63635	8	US-10-719-933-6874	Sequence 6874, Ap
565	13	65.0	3897	6	US-10-412-000-7	Sequence 7, Appli	c 638	13	65.0	70419	6	US-10-004-113-40	Sequence 40, Appl
566	13	65.0	4001	10	US-11-097-143-23569	Sequence 23569, A	c 639	13	65.0	71292	5	US-10-087-192-1942	Sequence 1942, Ap
567	13	65.0	4025	7	US-10-001-885-77	Sequence 77, Appl	c 640	13	65.0	75976	7	US-10-322-281-1942	Sequence 229, App
568	13	65.0	4025	10	US-11-057-447-77	Sequence 77, Appl	c 641	13	65.0	78313	7	US-10-052-482-124	Sequence 124, App
c 569	13	65.0	4026	5	US-10-171-277-17	Sequence 17, App	c 642	13	65.0	80246	3	US-09-728-552-4	Sequence 4, Appli
c 570	13	65.0	4026	5	US-10-981-287-115	Sequence 115, App	c 643	13	65.0	80595	3	US-09-728-552-3	Sequence 3, Appli
571	13	65.0	4169	9	US-10-915-740A-63	Sequence 63, Appl	c 644	13	65.0	84539	3	US-09-962-436-36	Sequence 36, Appl
c 572	13	65.0	4190	9	US-10-287-436A-186	Sequence 186, App	c 645	13	65.0	90442	5	US-10-105-637-1	Sequence 1, Appli
c 573	13	65.0	4190	9	US-10-287-436A-798	Sequence 798, App	c 646	13	65.0	90442	5	US-10-034-650-43	Sequence 43, Appl
c 574	13	65.0	4195	6	US-10-168-659-32	Sequence 32, Appl	c 647	13	65.0	90442	5	US-10-034-650-43	Sequence 43, Appl
c 575	13	65.0	4238	5	US-10-098-841-69	Sequence 69, Appl	c 648	13	65.0	92638	7	US-10-450-826-3	Sequence 3, Appli
576	13	65.0	4254	10	US-11-073-550-16	Sequence 16, Appl	c 649	13	65.0	94529	6	US-10-034-650-52	Sequence 52, Appl
577	13	65.0	4254	10	US-11-073-550-16	Sequence 16, Appl	c 650	13	65.0	94529	6	US-10-034-650-52	Sequence 52, Appl
c 578	13	65.0	4268	5	US-10-098-841-68	Sequence 68, Appl	c 651	13	65.0	100848	7	US-10-672-787-37	Sequence 37, Appl
c 579	13	65.0	4354	7	US-10-437-963-75382	Sequence 75382, A	c 652	13	65.0	100848	7	US-10-672-787-37	Sequence 37, Appl
c 580	13	65.0	4411	10	US-11-097-143-9235	Sequence 9235, Ap	c 653	13	65.0	103391	9	US-10-981-277-54	Sequence 54, Appl
c 581	13	65.0	4515	10	US-11-097-143-37387	Sequence 37387, A	c 654	13	65.0	133462	7	US-10-367-094-114	Sequence 114, App
582	13	65.0	4590	7	US-10-282-122A-3191	Sequence 3191, A	c 655	13	65.0	140040	7	US-10-275-762-69	Sequence 69, Appl
583	13	65.0	4616	7	US-10-398-221-3730	Sequence 3730, Ap	c 656	13	65.0	143601	3	US-09-855-824-3	Sequence 3, Appli
584	13	65.0	4826	2	US-08-781-986A-275	Sequence 275, App	c 657	13	65.0	143601	3	US-10-476-542-3	Sequence 3, Appli
585	13	65.0	4826	7	US-10-329-624-275	Sequence 275, App	c 658	13	65.0	146778	8	US-10-741-600-17710	Sequence 17710, A
c 586	13	65.0	4981	7	US-10-437-963-26537	Sequence 26537, A	c 659	13	65.0	155937	8	US-10-723-860-2208	Sequence 2208, Ap
c 587	13	65.0	5019	10	US-11-097-143-16408	Sequence 16408, A	c 660	13	65.0	161671	6	US-10-017-117-1	Sequence 1, Appli
c 588	13	65.0	5356	6	US-10-269-909-40	Sequence 40, Appl	c 661	13	65.0	172825	9	US-10-981-277-26	Sequence 26, Appl
c 589	13	65.0	5356	6	US-10-756-149-4427	Sequence 4427, Ap	c 662	13	65.0	177587	5	US-10-087-192-1438	Sequence 27, Appl
c 590	13	65.0	5358	6	US-10-264-049-944	Sequence 944, App	c 663	13	65.0	182508	9	US-10-981-277-27	Sequence 27, Appl
c 591	13	65.0	5621	10	US-11-097-143-40273	Sequence 40273, A	c 664	13	65.0	185053	8	US-10-417-375-125	Sequence 125, App
c 592	13	65.0	5813	9	US-10-956-157-4849	Sequence 4849, Ap	c 665	13	65.0	185053	8	US-10-322-281-800	Sequence 800, App
c 593	13	65.0	5817	8	US-10-723-860-1736	Sequence 1736, Ap	c 666	13	65.0	212231	5	US-10-087-192-1126	Sequence 1126, Ap
594	13	65.0	5853	8	US-10-723-860-6112	Sequence 6112, Ap	c 667	13	65.0	212231	5	US-10-087-192-1126	Sequence 1126, Ap
595	13	65.0	5972	10	US-11-097-143-6121	Sequence 6121, Ap	c 668	13	65.0	218802	9	US-10-473-939-10	Sequence 10, Appl
c 596	13	65.0	6069	6	US-10-085-198-27	Sequence 27, Appl	c 669	13	65.0	227246	5	US-10-897-508-1	Sequence 1, Appli
c 597	13	65.0	6157	6	US-10-085-198-29	Sequence 29, Appl	c 670	13	65.0	233060	5	US-10-322-281-314	Sequence 314, App
c 598	13	65.0	6195	6	US-10-085-198-25	Sequence 25, Appl	c 671	13	65.0	233060	5	US-10-087-192-97	Sequence 97, Appl
599	13	65.0	6254	3	US-09-860-670-279	Sequence 279, App	c 672	13	65.0	238484	5	US-10-087-192-544	Sequence 544, App
600	13	65.0	6254	6	US-10-074-024-674	Sequence 674, App	c 673	13	65.0	314364	9	US-10-917-647-3	Sequence 3, Appli
601	13	65.0	6254	6	US-10-227-646-279	Sequence 279, App	c 674	13	65.0	334462	9	US-10-496-011-1	Sequence 1, Appli
602	13	65.0	6261	6	US-10-437-963-49319	Sequence 49319, A	c 675	13	65.0	34462	9	US-10-322-696-76	Sequence 76, Appl
c 603	13	65.0	6360	6	US-10-223-538-9	Sequence 9, Appli	c 676	13	65.0	358246	6	US-10-292-798-1095	Sequence 1095, Ap
c 604	13	65.0	6835	10	US-11-097-143-9322	Sequence 9322, Ap	c 677	13	65.0	380963	9	US-10-737-082-5	Sequence 5, Appli
c 605	13	65.0	7018	6	US-10-062-674-1565	Sequence 1565, Ap	c 678	13	65.0	380963	9	US-10-765-790-5	Sequence 5, Appli
606	13	65.0	7159	2	US-08-781-986A-302	Sequence 302, App	c 679	13	65.0	383432	9	US-10-737-082-34	Sequence 34, Appl
607	13	65.0	7159	7	US-10-329-624-302	Sequence 302, App	c 680	13	65.0	383432	9	US-10-765-790-34	Sequence 34, Appl

c 681	13	65.0	546025	8	US-10-719-993-6862	Sequence 6862, Ap	c 754	12	60.0	163	7	US-10-282-122A-1617	Sequence 1617, Ap
c 682	13	65.0	684707	7	US-10-398-221-9	Sequence 9, Appli	c 755	12	60.0	184	8	US-10-425-115-165038	Sequence 165038,
c 683	13	65.0	721377	9	US-10-461-862-163	Sequence 163, App	c 756	12	60.0	189	7	US-10-437-963-75501	Sequence 75501, A
c 684	13	65.0	822900	6	US-10-292-798-1393	Sequence 1393, Ap	c 757	12	60.0	193	7	US-10-242-535A-47068	Sequence 47068, A
c 685	13	65.0	1691139	5	US-10-067-514-1	Sequence 1, Appli	c 758	12	60.0	193	7	US-10-282-122A-2707	Sequence 2707, Ap
c 686	13	65.0	1691139	7	US-10-419-723-1	Sequence 1, Appli	c 759	12	60.0	193	7	US-10-085-783A-47068	Sequence 47068, A
c 687	13	65.0	1691139	9	US-10-255-120-1	Sequence 1, Appli	c 760	12	60.0	196	7	US-10-424-599-33593	Sequence 33593, A
c 688	13	65.0	1691140	9	US-10-868-397-1	Sequence 1, Appli	c 761	12	60.0	201	7	US-10-741-601-19753	Sequence 19753, A
c 689	13	65.0	2242716	9	US-10-915-740A-1068	Sequence 1068, Ap	c 762	12	60.0	201	7	US-10-741-601-22371	Sequence 22371, A
c 690	13	65.0	2731748	7	US-10-397-465A-1	Sequence 1, Appli	c 763	12	60.0	201	8	US-10-719-993-13159	Sequence 13159, A
c 691	13	65.0	3011208	7	US-10-398-221-2058	Sequence 2058, Ap	c 764	12	60.0	201	8	US-10-719-993-16271	Sequence 16271, A
c 692	13	65.0	3186778	5	US-10-027-632-174961	Sequence 174961,	c 765	12	60.0	201	8	US-10-719-993-22969	Sequence 22969, A
c 693	13	65.0	3186778	6	US-10-027-632-174961	Sequence 174961,	c 766	12	60.0	201	8	US-10-719-993-24785	Sequence 24785, A
c 694	12	60.0	20	7	US-10-688-706-190	Sequence 190, App	c 767	12	60.0	201	8	US-10-719-993-24786	Sequence 24786, A
c 695	12	60.0	229	7	US-10-688-706-229	Sequence 229, App	c 768	12	60.0	201	8	US-10-719-993-44456	Sequence 44456, A
c 696	12	60.0	24	5	US-10-021-758-9	Sequence 9, Appli	c 769	12	60.0	201	8	US-10-741-600-34651	Sequence 34651, A
c 697	12	60.0	25	5	US-10-098-263B-83538	Sequence 83538, A	c 770	12	60.0	201	8	US-10-741-600-34726	Sequence 34726, A
c 698	12	60.0	25	7	US-10-719-956-1044	Sequence 1044, Ap	c 771	12	60.0	201	8	US-10-741-600-42802	Sequence 42802, A
c 699	12	60.0	25	7	US-10-719-956-37812	Sequence 37812, A	c 772	12	60.0	201	8	US-10-741-600-42802	Sequence 42802, A
c 700	12	60.0	25	7	US-10-719-956-150891	Sequence 150891,	c 773	12	60.0	201	8	US-10-741-600-53676	Sequence 53676, A
c 701	12	60.0	25	7	US-10-719-956-172010	Sequence 172010,	c 774	12	60.0	201	8	US-10-741-600-54497	Sequence 54497, A
c 702	12	60.0	25	7	US-10-719-956-182168	Sequence 182168,	c 775	12	60.0	209	5	US-10-741-600-62617	Sequence 62617, A
c 703	12	60.0	25	7	US-10-719-956-247146	Sequence 247146,	c 776	12	60.0	209	6	US-10-027-632-54865	Sequence 54865, A
c 704	12	60.0	25	7	US-10-719-956-247147	Sequence 247147,	c 777	12	60.0	211	3	US-10-425-115-161950	Sequence 161950,
c 705	12	60.0	25	7	US-10-719-956-519506	Sequence 519506,	c 778	12	60.0	212	3	US-09-764-860-74	Sequence 74, Appl
c 706	12	60.0	25	7	US-10-719-956-519506	Sequence 519506,	c 779	12	60.0	212	5	US-10-074-095-74	Sequence 74, Appl
c 707	12	60.0	25	7	US-10-719-956-546709	Sequence 546709,	c 780	12	60.0	212	6	US-10-212-872-74	Sequence 74, Appl
c 708	12	60.0	25	7	US-10-719-956-589609	Sequence 589609,	c 781	12	60.0	214	3	US-09-764-860-737	Sequence 737, App
c 709	12	60.0	25	7	US-10-719-956-618632	Sequence 618632,	c 782	12	60.0	214	3	US-09-764-860-738	Sequence 738, App
c 710	12	60.0	25	7	US-10-719-956-657240	Sequence 657240,	c 783	12	60.0	214	5	US-10-074-095-737	Sequence 737, App
c 711	12	60.0	25	8	US-10-719-900-133201	Sequence 133201,	c 784	12	60.0	214	5	US-10-074-095-738	Sequence 738, App
c 712	12	60.0	25	8	US-10-719-900-133202	Sequence 133202,	c 785	12	60.0	214	6	US-10-212-872-737	Sequence 737, App
c 713	12	60.0	25	8	US-10-719-900-178465	Sequence 178465,	c 786	12	60.0	214	6	US-10-212-872-738	Sequence 738, App
c 714	12	60.0	25	8	US-10-719-900-811681	Sequence 811681,	c 787	12	60.0	216	7	US-10-437-963-48443	Sequence 48443, A
c 715	12	60.0	25	8	US-10-719-900-826038	Sequence 826038,	c 788	12	60.0	218	8	US-10-674-124A-15647	Sequence 15647, A
c 716	12	60.0	25	8	US-10-719-900-898168	Sequence 898168,	c 789	12	60.0	218	8	US-10-674-124A-16096	Sequence 16096, A
c 717	12	60.0	25	8	US-10-719-900-932396	Sequence 932396,	c 790	12	60.0	220	7	US-10-437-963-71035	Sequence 71035, A
c 718	12	60.0	25	8	US-10-719-900-940488	Sequence 940488,	c 791	12	60.0	221	7	US-10-282-122A-1512	Sequence 1512, Ap
c 719	12	60.0	25	8	US-10-719-900-977047	Sequence 977047,	c 792	12	60.0	227	3	US-09-754-066-19	Sequence 19, Appl
c 720	12	60.0	25	10	US-11-036-317-538730	Sequence 538730,	c 793	12	60.0	227	7	US-10-768-089-19	Sequence 19, Appl
c 721	12	60.0	25	10	US-11-036-317-55062	Sequence 55062,	c 794	12	60.0	230	8	US-10-425-115-16185	Sequence 16185, A
c 722	12	60.0	25	10	US-11-036-317-583338	Sequence 583338,	c 795	12	60.0	232	6	US-10-062-674-42	Sequence 42, Appl
c 723	12	60.0	25	10	US-11-036-317-596996	Sequence 596996,	c 796	12	60.0	234	7	US-10-282-122A-2823	Sequence 2823, Ap
c 724	12	60.0	25	10	US-11-036-317-728903	Sequence 728903,	c 797	12	60.0	234	7	US-10-424-599-130533	Sequence 130533,
c 725	12	60.0	25	10	US-11-036-317-840190	Sequence 840190,	c 798	12	60.0	235	8	US-10-425-115-90389	Sequence 90389, A
c 726	12	60.0	25	10	US-11-036-317-864667	Sequence 864667,	c 799	12	60.0	239	6	US-10-247-813-24	Sequence 24, Appl
c 727	12	60.0	25	10	US-11-036-317-908508	Sequence 908508,	c 800	12	60.0	241	8	US-10-674-124A-15648	Sequence 15648, A
c 728	12	60.0	25	10	US-11-060-756-67199	Sequence 67199, A	c 801	12	60.0	245	7	US-10-627-476-649	Sequence 649, App
c 729	12	60.0	25	10	US-11-060-756-67200	Sequence 67200, A	c 802	12	60.0	245	8	US-10-674-124A-26568	Sequence 26568, A
c 730	12	60.0	25	10	US-11-060-756-67221	Sequence 67221, A	c 803	12	60.0	251	7	US-10-424-599-17575	Sequence 17575, A
c 731	12	60.0	25	10	US-11-060-756-67222	Sequence 67222, A	c 804	12	60.0	253	7	US-10-424-599-27488	Sequence 27488, A
c 732	12	60.0	25	10	US-11-060-756-131265	Sequence 131265,	c 805	12	60.0	253	8	US-10-425-115-39374	Sequence 39374, A
c 733	12	60.0	25	10	US-11-060-756-132091	Sequence 132091,	c 806	12	60.0	256	7	US-10-437-963-60791	Sequence 60791, A
c 734	12	60.0	25	10	US-11-060-756-132092	Sequence 132092,	c 807	12	60.0	258	7	US-10-424-599-637	Sequence 637, App
c 735	12	60.0	25	10	US-11-060-756-180424	Sequence 180424,	c 808	12	60.0	260	8	US-10-674-124A-7880	Sequence 7880, Ap
c 736	12	60.0	25	10	US-11-060-756-256701	Sequence 256701,	c 809	12	60.0	263	7	US-10-424-599-40241	Sequence 40241, A
c 737	12	60.0	25	10	US-11-060-756-283339	Sequence 283339,	c 810	12	60.0	266	8	US-10-425-115-118309	Sequence 118309,
c 738	12	60.0	25	10	US-11-060-756-287939	Sequence 287939,	c 811	12	60.0	270	7	US-10-242-535A-38942	Sequence 38942, A
c 739	12	60.0	25	10	US-11-060-756-296567	Sequence 296567,	c 812	12	60.0	270	7	US-10-424-599-36893	Sequence 36893, A
c 740	12	60.0	36	8	US-10-822-053-2	Sequence 2, Appli	c 813	12	60.0	270	7	US-10-085-783A-38942	Sequence 38942, A
c 741	12	60.0	36	8	US-10-822-053-18	Sequence 18, Appl	c 814	12	60.0	274	8	US-10-425-115-45151	Sequence 45151, A
c 742	12	60.0	40	3	US-09-829-113A-2	Sequence 2, Appli	c 815	12	60.0	280	3	US-09-732-627A-2804	Sequence 2804, Ap
c 743	12	60.0	41	7	US-10-035-833A-2647	Sequence 2647, Ap	c 816	12	60.0	280	7	US-10-282-122A-2455	Sequence 2455, Ap
c 744	12	60.0	41	7	US-10-035-833A-6866	Sequence 6866, Ap	c 817	12	60.0	281	6	US-10-029-386-16572	Sequence 16572, A
c 745	12	60.0	80	7	US-10-384-245-343	Sequence 343, App	c 818	12	60.0	283	8	US-10-357-930-50004	Sequence 50004, A
c 746	12	60.0	115	7	US-10-767-701-21818	Sequence 21818, A	c 819	12	60.0	286	8	US-10-425-115-157369	Sequence 157369,
c 747	12	60.0	119	3	US-09-770-696-695	Sequence 695, App	c 820	12	60.0	290	7	US-10-437-963-56467	Sequence 56467, A
c 748	12	60.0	121	6	US-10-307-005-181	Sequence 181, App	c 821	12	60.0	291	7	US-10-457-372-1	GENERAL INFORMA
c 749	12	60.0	121	6	US-10-307-005-182	Sequence 182, App	c 822	12	60.0	296	3	US-09-294-093B-2690	Sequence 2690, Ap
c 750	12	60.0	121	6	US-10-793-190-34	Sequence 34, Appl	c 823	12	60.0	296	8	US-10-425-115-38010	Sequence 38010, A
c 751	12	60.0	136	9	US-10-926-683-704	Sequence 704, App	c 824	12	60.0	296	8	US-10-425-115-68344	Sequence 68344, A
c 752	12	60.0	161	3	US-09-728-445-352	Sequence 352, App	c 825	12	60.0	296	8	US-10-425-115-126413	Sequence 126413,
c 753	12	60.0	161	9	US-10-964-549-352	Sequence 352, App	c 826	12	60.0	299	5	US-10-027-632-171575	Sequence 171575,

C 827	12	60.0	299	6	US-10-027-632-171575	Sequence 171575,	900	12	60.0	390	3	US-09-741-669-269	Sequence 269, App
C 828	12	60.0	300	3	US-09-764-891-8693	Sequence 8693, Ap	901	12	60.0	390	3	US-09-815-242-6239	Sequence 6239, Ap
C 829	12	60.0	300	7	US-10-424-599-126974	Sequence 126974,	902	12	60.0	390	3	US-09-815-242-9998	Sequence 9998, Ap
C 830	12	60.0	300	9	US-10-779-543-6856	Sequence 6856, Ap	903	12	60.0	390	6	US-10-287-274-216	Sequence 216, App
C 831	12	60.0	302	7	US-10-021-323-12118	Sequence 12118, A	904	12	60.0	390	6	US-10-287-274-248	Sequence 248, App
C 832	12	60.0	304	3	US-09-783-900-3935	Sequence 3935, Ap	905	12	60.0	390	7	US-10-282-122A-6591	Sequence 6591, Ap
C 833	12	60.0	304	3	US-09-864-408A-1483	Sequence 1483, Ap	906	12	60.0	390	7	US-10-282-122A-39550	Sequence 39550, A
C 834	12	60.0	304	7	US-10-424-599-103300	Sequence 103300, A	907	12	60.0	390	7	US-10-282-122A-42099	Sequence 42099, A
C 835	12	60.0	306	7	US-10-021-323-12206	Sequence 12206, A	C 908	12	60.0	394	7	US-10-424-599-27090	Sequence 27090, A
C 836	12	60.0	309	7	US-10-424-599-108677	Sequence 108677,	C 909	12	60.0	394	7	US-10-424-599-47124	Sequence 47124, A
C 837	12	60.0	316	7	US-10-424-599-18502	Sequence 18502, A	C 910	12	60.0	394	7	US-10-437-963-35094	Sequence 35094, A
C 838	12	60.0	324	7	US-10-424-599-62561	Sequence 62561, A	C 911	12	60.0	395	3	US-09-764-877-341	Sequence 341, App
C 839	12	60.0	324	3	US-10-437-963-63043	Sequence 63043, A	C 912	12	60.0	395	3	US-09-918-995-18376	Sequence 18376, A
C 840	12	60.0	326	3	US-09-796-692-3485	Sequence 3485, Ap	C 913	12	60.0	395	3	US-09-732-627A-1739	Sequence 1739, Ap
C 841	12	60.0	326	5	US-10-040-862-3485	Sequence 3485, Ap	C 914	12	60.0	395	6	US-10-242-515-941	Sequence 341, App
C 842	12	60.0	326	6	US-10-057-475B-3485	Sequence 3485, Ap	C 915	12	60.0	395	8	US-10-357-930-10059	Sequence 10059, A
C 843	12	60.0	326	6	US-10-154-884B-3485	Sequence 3485, Ap	C 916	12	60.0	397	7	US-10-424-599-17694	Sequence 17694, A
C 844	12	60.0	326	8	US-10-764-324-3485	Sequence 3485, Ap	C 917	12	60.0	397	8	US-10-425-115-95057	Sequence 95057, A
C 845	12	60.0	328	8	US-10-425-115-78137	Sequence 78137, A	C 918	12	60.0	398	3	US-09-960-706-79	Sequence 79, Appl
C 846	12	60.0	329	7	US-10-242-535A-33051	Sequence 33051, A	C 919	12	60.0	398	3	US-09-873-319-54	Sequence 54, Appl
C 847	12	60.0	329	7	US-10-085-783A-33051	Sequence 33051, A	C 920	12	60.0	399	7	US-10-282-122A-2683	Sequence 2683, Ap
C 848	12	60.0	330	8	US-10-425-115-140772	Sequence 140772,	C 921	12	60.0	399	8	US-10-425-115-162762	Sequence 162762,
C 849	12	60.0	331	7	US-10-242-535A-42416	Sequence 42416, A	C 922	12	60.0	402	9	US-10-450-763-14285	Sequence 14285, A
C 850	12	60.0	331	7	US-10-085-783A-42416	Sequence 42416, A	C 923	12	60.0	405	8	US-10-425-115-28016	Sequence 28016, A
C 851	12	60.0	332	8	US-10-425-115-122441	Sequence 122441,	C 924	12	60.0	406	7	US-10-282-122A-3198	Sequence 3198, Ap
C 852	12	60.0	333	7	US-10-437-963-87423	Sequence 87423, A	C 925	12	60.0	406	9	US-10-756-149-2743	Sequence 2743, Ap
C 853	12	60.0	334	8	US-10-425-115-33079	Sequence 33079, A	C 926	12	60.0	411	7	US-10-424-599-69933	Sequence 69933, A
C 854	12	60.0	336	8	US-10-424-599-48923	Sequence 48923, A	C 927	12	60.0	412	7	US-10-437-963-91069	Sequence 91069, A
C 855	12	60.0	336	8	US-10-425-115-11658	Sequence 11658, A	C 928	12	60.0	417	7	US-10-434-599-116469	Sequence 116469,
C 856	12	60.0	337	5	US-10-027-632-56700	Sequence 56700, A	C 929	12	60.0	420	3	US-09-922-293-9	Sequence 9, Appl
C 857	12	60.0	337	5	US-10-027-632-302195	Sequence 302195,	C 930	12	60.0	420	8	US-10-723-860-2817	Sequence 2817, Ap
C 858	12	60.0	337	6	US-10-027-632-56700	Sequence 56700, A	C 931	12	60.0	420	9	US-10-972-079-19084	Sequence 19084, A
C 859	12	60.0	337	6	US-10-027-632-302195	Sequence 302195,	C 932	12	60.0	422	8	US-10-357-930-17973	Sequence 17973, A
C 860	12	60.0	342	8	US-10-425-115-59319	Sequence 59319, A	C 933	12	60.0	422	8	US-09-925-065A-555287	Sequence 555287,
C 861	12	60.0	343	8	US-10-425-115-118006	Sequence 118006,	C 934	12	60.0	422	9	US-10-767-701-21547	Sequence 21547, A
C 862	12	60.0	343	8	US-10-653-047-3747	Sequence 3747, Ap	C 935	12	60.0	422	9	US-10-450-763-17703	Sequence 17703, A
C 863	12	60.0	345	3	US-09-783-590-3871	Sequence 3871, Ap	C 936	12	60.0	423	3	US-09-974-300-7094	Sequence 7094, Ap
C 864	12	60.0	345	3	US-10-424-599-9225	Sequence 9225, Ap	C 937	12	60.0	423	7	US-10-437-963-2993	Sequence 2993, Ap
C 865	12	60.0	346	8	US-10-425-115-92012	Sequence 92012, A	C 938	12	60.0	428	7	US-10-424-599-15564	Sequence 15564, A
C 866	12	60.0	350	7	US-10-424-599-32414	Sequence 32414, A	C 939	12	60.0	429	8	US-10-424-599-17430	Sequence 17430, A
C 867	12	60.0	350	7	US-10-437-963-93907	Sequence 93907, A	C 940	12	60.0	429	8	US-10-425-115-168784	Sequence 168784,
C 868	12	60.0	351	4	US-09-925-065A-596207	Sequence 596207, A	C 941	12	60.0	432	4	US-09-925-065A-390569	Sequence 390569,
C 869	12	60.0	351	8	US-10-425-115-33082	Sequence 33082, A	C 942	12	60.0	433	8	US-10-425-115-114755	Sequence 114755,
C 870	12	60.0	353	7	US-10-424-599-104851	Sequence 104851,	C 943	12	60.0	435	8	US-10-425-115-73748	Sequence 73748, A
C 871	12	60.0	353	7	US-10-609-021-2	Sequence 2, Appli	C 944	12	60.0	435	8	US-10-425-115-162015	Sequence 162015,
C 872	12	60.0	353	8	US-10-425-115-134992	Sequence 134992,	C 945	12	60.0	437	4	US-09-925-065A-315234	Sequence 315234,
C 873	12	60.0	353	9	US-10-779-543-15668	Sequence 15668, A	C 946	12	60.0	438	7	US-10-282-122A-2350	Sequence 2350, Ap
C 874	12	60.0	354	8	US-10-357-930-53890	Sequence 53890, A	C 947	12	60.0	438	7	US-10-282-122A-2350	Sequence 2350, Ap
C 875	12	60.0	357	4	US-09-925-065A-16043	Sequence 16043, A	C 948	12	60.0	438	7	US-10-424-599-55166	Sequence 55166, A
C 876	12	60.0	362	7	US-10-424-599-70268	Sequence 70268, A	C 949	12	60.0	439	3	US-09-983-965-3675	Sequence 3675, Ap
C 877	12	60.0	369	8	US-10-425-115-152634	Sequence 152634,	C 950	12	60.0	439	5	US-10-052-283-145	Sequence 145, Ap
C 878	12	60.0	373	8	US-10-425-115-174056	Sequence 174056,	C 951	12	60.0	440	3	US-09-783-590-10547	Sequence 10547, A
C 879	12	60.0	374	3	US-09-886-242A-3	Sequence 3, Appli	C 952	12	60.0	440	7	US-10-424-599-2813	Sequence 2813, Ap
C 880	12	60.0	374	5	US-10-027-603-3	Sequence 3, Appli	C 953	12	60.0	440	7	US-10-437-963-18378	Sequence 18378, A
C 881	12	60.0	374	8	US-10-425-115-118446	Sequence 118446,	C 954	12	60.0	442	5	US-10-027-632-312948	Sequence 312948,
C 882	12	60.0	374	9	US-10-692-299-3	Sequence 3, Appli	C 955	12	60.0	442	6	US-10-027-632-312948	Sequence 312948,
C 883	12	60.0	375	7	US-10-424-599-17132	Sequence 17132, A	C 956	12	60.0	442	6	US-10-424-599-93156	Sequence 93156, A
C 884	12	60.0	376	7	US-10-282-122A-2794	Sequence 2794, Ap	C 957	12	60.0	445	4	US-09-925-065A-468946	Sequence 468946,
C 885	12	60.0	376	8	US-10-425-115-165553	Sequence 165553,	C 958	12	60.0	446	4	US-10-369-493-33388	Sequence 33388, A
C 886	12	60.0	377	8	US-10-425-115-74830	Sequence 74830, A	C 959	12	60.0	447	6	US-10-425-114-34360	Sequence 34360, A
C 887	12	60.0	380	3	US-09-732-627A-3995	Sequence 3995, Ap	C 960	12	60.0	447	7	US-09-878-178-544	Sequence 544, App
C 888	12	60.0	380	8	US-10-357-930-15127	Sequence 15127, A	C 961	12	60.0	448	5	US-10-046-935-544	Sequence 544, App
C 889	12	60.0	385	8	US-10-425-115-158050	Sequence 158050,	C 962	12	60.0	448	5	US-10-146-502-544	Sequence 544, App
C 890	12	60.0	386	7	US-10-424-599-117649	Sequence 117649,	C 963	12	60.0	449	4	US-09-925-065A-500670	Sequence 500670,
C 891	12	60.0	387	3	US-09-983-965-2883	Sequence 2883, Ap	C 964	12	60.0	449	4	US-09-925-065A-500671	Sequence 500671,
C 892	12	60.0	387	7	US-10-282-122A-19945	Sequence 19945, A	C 965	12	60.0	450	8	US-10-425-115-130230	Sequence 130230,
C 893	12	60.0	387	7	US-10-282-122A-23146	Sequence 23146, A	C 966	12	60.0	451	4	US-09-925-065A-536155	Sequence 536155,
C 894	12	60.0	387	7	US-10-282-122A-37371	Sequence 37371, A	C 967	12	60.0	451	4	US-09-925-065A-555285	Sequence 555285,
C 895	12	60.0	387	7	US-10-767-701-26639	Sequence 26639, A	C 968	12	60.0	451	4	US-09-925-065A-555286	Sequence 555286,
C 896	12	60.0	388	3	US-09-878-178-843	Sequence 843, App	C 969	12	60.0	452	4	US-09-925-065A-267197	Sequence 267197,
C 897	12	60.0	388	5	US-10-046-935-843	Sequence 843, App	C 970	12	60.0	452	4	US-09-925-065A-267198	Sequence 267198,
C 898	12	60.0	388	5	US-10-146-502-843	Sequence 843, App	C 971	12	60.0	455	4	US-09-925-065A-269825	Sequence 269825,
C 899	12	60.0	388	7	US-10-424-599-90599	Sequence 90599, A	C 972	12	60.0	455	8	US-10-425-115-69754	Sequence 69754, A

```
c 973      12 60.0 457 4 US-09-925-065A-471320
c 974      12 60.0 457 4 US-09-925-065A-471321
c 975      12 60.0 457 4 US-09-925-065A-471322
c 976      12 60.0 457 7 US-10-424-599-8176
c 977      12 60.0 458 5 US-10-027-632-55485
c 978      12 60.0 458 5 US-10-027-632-36257
c 979      12 60.0 458 5 US-10-027-632-302792
c 980      12 60.0 458 5 US-10-027-632-302793
c 981      12 60.0 458 6 US-10-027-632-55485
c 982      12 60.0 458 6 US-10-027-632-56257
c 983      12 60.0 458 6 US-10-027-632-302792
c 984      12 60.0 458 6 US-10-027-632-302793
c 985      12 60.0 459 4 US-09-925-065A-621852
c 986      12 60.0 459 8 US-10-357-930-33043
c 987      12 60.0 461 5 US-10-102-524-1191
c 988      12 60.0 461 7 US-10-437-963-81908
c 989      12 60.0 462 5 US-10-027-632-2915
c 990      12 60.0 462 5 US-10-027-632-36847
c 991      12 60.0 462 6 US-10-027-632-2915
c 992      12 60.0 462 6 US-10-027-632-36847
c 993      12 60.0 467 3 US-09-764-870-56
c 994      12 60.0 467 3 US-09-764-853-216
c 995      12 60.0 467 5 US-10-125-540-56
c 996      12 60.0 467 5 US-10-103-313-142
c 997      12 60.0 467 7 US-10-158-057-84
c 998      12 60.0 467 8 US-10-674-124A-13551
c 999      12 60.0 468 4 US-09-925-065A-434070
c1000     12 60.0 470 3 US-09-864-761-14460
```

## ALIGNMENTS

```
RESULT 1
US-10-671-074-172
; Sequence 172, Application US/10671074
; Publication No. US20040097459A1
; GENERAL INFORMATION:
; APPLICANT: Dobie, Kenneth W.
; APPLICANT: Bhanot, Sanjay
; APPLICANT: Veniat-Elisson, Murielle
; APPLICANT: Lindberg, Richard A.
; APPLICANT: Shutter, John R.
; TITLE OF INVENTION: MODULATION OF FORKHEAD BOX O1A EXPRESSION
; FILE REFERENCE: AMGN0001-101
; CURRENT APPLICATION NUMBER: US/10/671,074
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 10/260,203
; PRIOR FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 172
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-671-074-172
```

```
Query Match      100.0%; Score 20; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GCTTTGGTTGGGCAACACAT 20
      |||||||
Db      1 GCTTTGGTTGGGCAACACAT 20
```

```
RESULT 2
US-10-719-900-85415/c
; Sequence 85415, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
```

```
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 85415
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-85415
```

```
Query Match      100.0%; Score 20; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GCTTTGGTTGGGCAACACAT 20
      |||||||
Db      21 GCTTTGGTTGGGCAACACAT 2
```

```
RESULT 3
US-09-794-928A-12/c
; Sequence 12, Application US/09794928A
; Patent No. US20020127555A1
; GENERAL INFORMATION:
; APPLICANT: Baban, Soheyl
; APPLICANT: Bernard, Monique
; APPLICANT: Cherry, Elana
; APPLICANT: Gosselin, Diane
; APPLICANT: Hugo, Patrice
; APPLICANT: Malette, Brigitte
; APPLICANT: Miron, Pierre
; APPLICANT: Prive, Charles
; APPLICANT: Shazand, Kamran
; TITLE OF INVENTION: ENDOMETRIOSIS-RELATED MARKERS AND USES THEREOF
; FILE REFERENCE: 5600-71
; CURRENT APPLICATION NUMBER: US/09/794,928A
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 60/225,745
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 60/185,063
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 254
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-794-928A-12
```

```
Query Match      100.0%; Score 20; DB 3; Length 254;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GCTTTGGTTGGGCAACACAT 20
      |||||||
Db      90 GCTTTGGTTGGGCAACACAT 71
```

```
RESULT 4
US-10-242-535A-24143/c
; Sequence 24143, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
```

```

RESULT 5
US-10-085-783A-24143/c
; Sequence 24143, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: Chondrogene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24143

```

```

RESULT 6
US-10-242-535A-35649/c
; Sequence 35649, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C. C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35649
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-35649

```



```
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTTTGGTTGGCAACACAT 20
Db 2647 GCTTTGGTTGGCAACACAT 2628

RESULT 11
US-10-391-530-3/c
; Sequence 3, Application US/10391530
; Publication No. US20040110227A1
; GENERAL INFORMATION:
; APPLICANT: Levanon, Erez
; APPLICANT: Toporik, Amir
; APPLICANT: Akiva, Pini
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR IDENTIFYING PUTATIVE FUSION TRANSCRIPTS,
; TITLE OF INVENTION: POLYPEPTIDES ENCODED THEREFROM AND POLYNUCLEOTIDE SEQUENCES RELA
; FILE REFERENCE: 25835
; CURRENT APPLICATION NUMBER: US/10/391,530
; CURRENT FILING DATE: 2003-03-19
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 3200
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-391-530-3
Query Match 100.0%; Score 20; DB 7; Length 3200;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTTTGGTTGGCAACACAT 20
Db 2511 GCTTTGGTTGGCAACACAT 2492

RESULT 12
US-10-391-530-2/c
; Sequence 2, Application US/10391530
; Publication No. US20040110227A1
; GENERAL INFORMATION:
; APPLICANT: Levanon, Erez
; APPLICANT: Toporik, Amir
; APPLICANT: Akiva, Pini
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR IDENTIFYING PUTATIVE FUSION TRANSCRIPTS,
; TITLE OF INVENTION: POLYPEPTIDES ENCODED THEREFROM AND POLYNUCLEOTIDE SEQUENCES RELA
; FILE REFERENCE: 25835
; CURRENT APPLICATION NUMBER: US/10/391,530
; CURRENT FILING DATE: 2003-03-19
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 3517
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-391-530-2
Query Match 100.0%; Score 20; DB 7; Length 3517;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTTTGGTTGGCAACACAT 20
Db 2022 GCTTTGGTTGGCAACACAT 2003

RESULT 13
US-10-671-074-11/c
; Sequence 11, Application US/10671074
; Publication No. US20040097459A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Dobie, Kenneth W.
; APPLICANT: Bhanot, Sanjay
; APPLICANT: Veniant-Ellison, Murielle
; APPLICANT: Lindberg, Richard A.
; APPLICANT: Shutter, John R.
; TITLE OF INVENTION: MODULATION OF FORKHEAD BOX OIA EXPRESSION
; FILE REFERENCE: AMGN0001-101
; CURRENT APPLICATION NUMBER: US/10/671,074
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 10/260,203
; PRIOR FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 11
; LENGTH: 4945
; TYPE: DNA
; ORGANISM: M. musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (429)...(2387)
US-10-671-074-11
Query Match 100.0%; Score 20; DB 7; Length 4945;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTTTGGTTGGCAACACAT 20
Db 2339 GCTTTGGTTGGCAACACAT 2320

RESULT 14
US-10-007-926A-134/c
; Sequence 134, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; FILE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 134
; LENGTH: 5723
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: forkhead box oia (rhabdomyosarcoma) (FOXO1A)
; OTHER INFORMATION: gene.
US-10-007-926A-134
Query Match 100.0%; Score 20; DB 6; Length 5723;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTTTGGTTGGCAACACAT 20
Db 2305 GCTTTGGTTGGCAACACAT 2286

RESULT 15
US-10-341-434-52/c
; Sequence 52, Application US/10341434
```

```
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 5723
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (386)..(2350)
; OTHER INFORMATION:
US-10-341-434-52

Query Match      100.0%; Score 20; DB 6; Length 5723;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCTTTGGTTGGCAACACAT 20
Db      2305 GCTTTGGTTGGCAACACAT 2286

RESULT 16
US-10-671-074-4/c
; Sequence 4, Application US/10671074
; Publication No. US20040097459A1
; GENERAL INFORMATION:
; APPLICANT: Dobie, Kenneth W.
; APPLICANT: Bhanot, Sanjay
; APPLICANT: Veniant-Ellison, Murielle
; APPLICANT: Lindberg, Richard A.
; APPLICANT: Shutter, John R.
; TITLE OF INVENTION: MODULATION OF FORKHEAD BOX O1A EXPRESSION
; FILE REFERENCE: AMGN0001-101
; CURRENT APPLICATION NUMBER: US/10/671,074
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 10/260,203
; PRIOR FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 4
; LENGTH: 5723
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (386)...(2353)
US-10-671-074-4

Query Match      100.0%; Score 20; DB 7; Length 5723;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCTTTGGTTGGCAACACAT 20
Db      2305 GCTTTGGTTGGCAACACAT 2286

RESULT 17
US-10-956-157-760/c
; Sequence 760, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 760
; LENGTH: 5723
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-760

Query Match      100.0%; Score 20; DB 9; Length 5723;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCTTTGGTTGGCAACACAT 20
Db      2305 GCTTTGGTTGGCAACACAT 2286

RESULT 18
US-10-450-763-20857/c
; Sequence 20857, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 03/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 20857
; LENGTH: 5833
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (157)..(585)
; OTHER INFORMATION: 31% homologous to Homo sapiens fork head domain
; OTHER INFORMATION: protein, accession number U02310, Smith-Waterman Score=53.
US-10-450-763-20857

Query Match      100.0%; Score 20; DB 9; Length 5833;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCTTTGGTTGGCAACACAT 20
Db      2357 GCTTTGGTTGGCAACACAT 2338

RESULT 19
US-10-671-074-40
; Sequence 40, Application US/10671074
; Publication No. US20040097459A1
; GENERAL INFORMATION:
; APPLICANT: Dobie, Kenneth W.
; APPLICANT: Bhanot, Sanjay
; APPLICANT: Veniant-Ellison, Murielle
; APPLICANT: Lindberg, Richard A.
; APPLICANT: Shutter, John R.
; TITLE OF INVENTION: MODULATION OF FORKHEAD BOX O1A EXPRESSION
```



; FILE REFERENCE: AMGN0001-101  
; CURRENT APPLICATION NUMBER: US/10/671,074  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: US 10/260,203  
; PRIOR FILING DATE: 2002-09-26  
; NUMBER OF SEQ ID NOS: 176  
; SEQ ID NO 40  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Antisense Oligonucleotide  
US-10-671-074-40

Query Match 95.0%; Score 19; DB 7; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTTTGGTTGGCAACACAT 20  
|||||  
DB 1 CTTTGGTTGGCAACACAT 19

RESULT 20  
US-10-671-074-118/c  
; Sequence 118, Application US/10671074  
; Publication No. US20040097459A1  
; GENERAL INFORMATION:  
; APPLICANT: Dobie, Kenneth W.  
; APPLICANT: Bhanot, Sanjay  
; APPLICANT: Veniant-Ellison, Murielle  
; APPLICANT: Lindberg, Richard A.  
; APPLICANT: Shutter, John R.  
; TITLE OF INVENTION: MODULATION OF FORKHEAD BOX O1A EXPRESSION  
; FILE REFERENCE: AMGN0001-101  
; CURRENT APPLICATION NUMBER: US/10/671,074  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: US 10/260,203  
; PRIOR FILING DATE: 2002-09-26  
; NUMBER OF SEQ ID NOS: 176  
; SEQ ID NO 118  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: H. sapiens  
; FEATURE:  
US-10-671-074-118

Query Match 95.0%; Score 19; DB 7; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTTTGGTTGGCAACACAT 20  
|||||  
DB 20 CTTTGGTTGGCAACACAT 2

RESULT 21  
US-09-938-842A-4239/c  
; Sequence 4239, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIP1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 4239  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-4239

Query Match 85.0%; Score 17; DB 3; Length 2000;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACA 17  
|||||  
DB 327 GCTTTGGTTGGCAACA 311

RESULT 22  
US-09-938-842A-4239/c  
; Sequence 4239, Application US/09938842A  
; Publication No. US2004009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIP1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 4239  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-4239

Query Match 85.0%; Score 17; DB 3; Length 2000;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAACA 17  
|||||  
DB 327 GCTTTGGTTGGCAACA 311

RESULT 23  
US-10-671-074-41  
; Sequence 41, Application US/10671074  
; Publication No. US20040097459A1  
; GENERAL INFORMATION:  
; APPLICANT: Dobie, Kenneth W.  
; APPLICANT: Bhanot, Sanjay  
; APPLICANT: Veniant-Ellison, Murielle  
; APPLICANT: Lindberg, Richard A.  
; APPLICANT: Shutter, John R.  
; TITLE OF INVENTION: MODULATION OF FORKHEAD BOX O1A EXPRESSION  
; FILE REFERENCE: AMGN0001-101  
; CURRENT APPLICATION NUMBER: US/10/671,074  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: US 10/260,203  
; PRIOR FILING DATE: 2002-09-26  
; NUMBER OF SEQ ID NOS: 176  
; SEQ ID NO 41



```

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 107214
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(405)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6782C.1
US-10-424-599-107214

Query Match          75.0%; Score 15; DB 7; Length 405;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 GGTGGGCAACACAT 20
        |||||
Db      277 GGTGGGCAACACAT 263

RESULT 29
US-10-430-201-4459
; Sequence 4459, Application US/10430201
; Publication No. US20040162679A1
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; TITLE OF INVENTION: Method for Predicting Gene Potential and Cell Commitment
; FILE REFERENCE: 40716 (IP-010)
; CURRENT APPLICATION NUMBER: US/10/430,201
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US 60/370,114
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 4879
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4459
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (180)..(180)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (185)..(185)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (226)..(226)
; OTHER INFORMATION: n is a, c, g, or t
US-10-430-201-4459

Query Match          75.0%; Score 15; DB 7; Length 406;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 GGTGGGCAACACAT 20
        |||||
Db      267 GGTGGGCAACACAT 281

RESULT 30
US-10-430-201-4460
; Sequence 4460, Application US/10430201
; Publication No. US20040162679A1
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; TITLE OF INVENTION: Method for Predicting Gene Potential and Cell Commitment

```

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA125  
; CURRENT APPLICATION NUMBER: US/09/764,872  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 957  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 161  
; LENGTH: 574  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-872-161

Query Match 75.0%; Score 15; DB 3; Length 574;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTTGGTTGGGCAACA 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 539 TTTGGTTGGGCAACA 553

RESULT 33  
US-10-425-115-78789/c  
; Sequence 78789, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 78789  
; LENGTH: 602  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_171882C.1  
US-10-425-115-78789

Query Match 75.0%; Score 15; DB 8; Length 602;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTTGGTTGGGCAACA 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 579 TTTGGTTGGGCAACA 565

RESULT 34  
US-09-925-065A-669805  
; Sequence 669805, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766

; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 669805  
; LENGTH: 1153  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-669805

Query Match 75.0%; Score 15; DB 4; Length 1153;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTTGGTTGGGCAAC 16  
| | | | | | | | | | | | | | | | | | | | | |  
Db 119 CTTTGGTTGGGCAAC 133

RESULT 35  
US-10-450-763-11672/c  
; Sequence 11672, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 11672  
; LENGTH: 1263  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (977) ..(2311)  
; OTHER INFORMATION: 88% homologous to Homo sapiens putative pl50, accession number  
; OTHER INFORMATION: U93574, Smith-Waterman Score=2196.  
US-10-450-763-11672

Query Match 75.0%; Score 15; DB 9; Length 1263;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTTTGGTTGGGCAA 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 255 GCTTTGGTTGGGCAA 241

RESULT 36  
US-09-956-004-66  
; Sequence 66, Application US/09956004  
; Patent No. US20020072595A1  
; GENERAL INFORMATION:  
; APPLICANT: Patrick J. Dillon et al.  
; TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands  
; FILE REFERENCE: PB324D1  
; CURRENT APPLICATION NUMBER: US/09/956,004  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: 08/976,259  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/061,953  
; PRIOR FILING DATE: 1997-10-14  
; PRIOR APPLICATION NUMBER: 60/031,626

; PRIOR FILING DATE: 1996-11-22  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 66  
; LENGTH: 1342  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1238)..(1238)  
; OTHER INFORMATION: n equals a, t, g, or c  
US-09-956-004-66

Query Match 75.0%; Score 15; DB 3; Length 1342;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAA 15  
|||  
Db 614 GCTTTGGTTGGCAA 628

## RESULT 37

US-10-808-570-66  
; Sequence 66, Application US/10808570  
; Publication No. US20040192903A1  
; GENERAL INFORMATION:  
; APPLICANT: Patrick J. Dillon et al.  
; TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands  
; FILE REFERENCE: PB324D1  
; CURRENT APPLICATION NUMBER: US/10/808,570  
; PRIOR FILING DATE: 2004-03-25  
; PRIOR APPLICATION NUMBER: US/09/956,004  
; PRIOR FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: 08/976,259  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/061,953  
; PRIOR FILING DATE: 1997-10-14  
; PRIOR APPLICATION NUMBER: 60/031,626  
; PRIOR FILING DATE: 1996-11-22  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 66  
; LENGTH: 1342  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1238)..(1238)  
; OTHER INFORMATION: n equals a, t, g, or c  
US-10-808-570-66

Query Match 75.0%; Score 15; DB 8; Length 1342;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAA 15  
|||  
Db 614 GCTTTGGTTGGCAA 628

## RESULT 38

US-10-108-260A-1546/c  
; Sequence 1546, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1546  
; LENGTH: 1550  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-108-260A-1546

Query Match 75.0%; Score 15; DB 6; Length 1550;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAA 15  
|||  
Db 269 GCTTTGGTTGGCAA 255

## RESULT 39

US-10-282-122A-9084  
; Sequence 9084, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9084  
; LENGTH: 2154  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-10-282-122A-9084

Query Match 75.0%; Score 15; DB 7; Length 2154;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTTGGTTGGCAA 15  
|||  
Db 860 GCTTTGGTTGGCAA 874

RESULT 40  
US-10-094-749-916  
; Sequence 916, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AJ  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOKYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 916  
; LENGTH: 3059  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-094-749-916

Query Match 75.0%; Score 15; DB 6; Length 3059;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20  
Db 1579 GGTGGGCAACACAT 1593

RESULT 41  
US-11-097-143-32554/c  
; Sequence 32554, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 916  
; LENGTH: 3366  
; TYPE: DNA  
; ORGANISM: DROSOPHILA  
US-11-097-143-32554

Query Match 75.0%; Score 15; DB 6; Length 3059;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20  
Db 1579 GGTGGGCAACACAT 1593

; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 32554  
; LENGTH: 3342  
; TYPE: DNA  
; ORGANISM: DROSOPHILA  
US-11-097-143-32554

Query Match 75.0%; Score 15; DB 10; Length 3342;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTTGGTTGGCAAC 16  
Db 648 CTTTGGTTGGCAAC 634

RESULT 42  
US-11-097-143-23078/c  
; Sequence 23078, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23078  
; LENGTH: 3366  
; TYPE: DNA  
; ORGANISM: DROSOPHILA  
US-11-097-143-23078

Query Match 75.0%; Score 15; DB 10; Length 3366;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CTTTGGTTGGCAAC 16  
Db 3075 CTTTGGTTGGCAAC 3061

RESULT 43  
US-10-002-775-10  
; Sequence 10, Application US/10002775  
; Publication No. US20020102651A1  
; GENERAL INFORMATION:

APPLICANT: Gordon Freeman  
APPLICANT: Vassiliki Bousiouris  
APPLICANT: Tatyana Chernova  
APPLICANT: Nelly Malenkovich  
TITLE OF INVENTION: NOVEL B7-4 MOLECULES AND USES THEREFOR  
FILE REFERENCE: GNN-004ADV  
CURRENT APPLICATION NUMBER: US/10/002,775  
CURRENT FILING DATE: 2001-11-02  
PRIOR APPLICATION NUMBER: US 09/644,934  
PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: 60/150,390  
PRIOR FILING DATE: 1999-08-23  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 3593  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (17)..(889)  
US-10-002-775-10

Query Match 75.0%; Score 15; DB 5; Length 3593;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GGTGGGCAACACAT 20  
|||||  
DB 3406 GGTGGGCAACACAT 3420

RESULT 44  
US-10-115-615-22  
Sequence 22, Application US/10115615  
Publication No. US20030044768A1  
GENERAL INFORMATION:  
APPLICANT: Wood, Clive  
APPLICANT: Freeman, Gordon  
APPLICANT: Chaudhary, Divya  
TITLE OF INVENTION: PD-1, A RECEPTOR FOR B7-4, AND USES THEREFOR  
FILE REFERENCE: GNN-004C  
CURRENT APPLICATION NUMBER: US/10/115,615  
CURRENT FILING DATE: 2002-04-02  
PRIOR APPLICATION NUMBER: US 60/281064  
PRIOR FILING DATE: 2001-04-02  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 22  
LENGTH: 3593  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (17)..(889)  
US-10-115-615-22

Query Match 75.0%; Score 15; DB 5; Length 3593;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GGTGGGCAACACAT 20  
|||||  
DB 3406 GGTGGGCAACACAT 3420

RESULT 45  
US-10-764-420-1546  
Sequence 1546, Application US/10764420  
Publication No. US20050084872A1  
GENERAL INFORMATION:  
APPLICANT: Lum, Pek Yee  
APPLICANT: Tan, Yejun

APPLICANT: Dai, Hongyue  
TITLE OF INVENTION: Methods For Determining Whether An Agent Possesses A Defined Biological Activity  
FILE REFERENCE: ROSA122057  
CURRENT APPLICATION NUMBER: US/10/764,420  
CURRENT FILING DATE: 2004-01-23  
PRIOR APPLICATION NUMBER: US 60/442,797  
PRIOR FILING DATE: 2003-01-24  
PRIOR APPLICATION NUMBER: US 60/474,413  
PRIOR FILING DATE: 2003-05-30  
NUMBER OF SEQ ID NOS: 3683  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1546  
LENGTH: 3593  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-764-420-1546

Query Match 75.0%; Score 15; DB 9; Length 3593;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GGTGGGCAACACAT 20  
|||||  
DB 3406 GGTGGGCAACACAT 3420

RESULT 46  
US-10-631-467-1140  
Sequence 1140, Application US/10631467  
Publication No. US20050208496A1  
GENERAL INFORMATION:  
APPLICANT: Genox Research Inc.  
TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive  
FILE REFERENCE: 3462.1005-000  
CURRENT APPLICATION NUMBER: US/10/631,467  
CURRENT FILING DATE: 2003-07-31  
PRIOR APPLICATION NUMBER: JP 2003-077212  
PRIOR FILING DATE: 2003-03-20  
PRIOR APPLICATION NUMBER: JP 2002-229312  
PRIOR FILING DATE: 2002-08-06  
NUMBER OF SEQ ID NOS: 2086  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1140  
LENGTH: 3593  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-631-467-1140

Query Match 75.0%; Score 15; DB 9; Length 3593;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GGTGGGCAACACAT 20  
|||||  
DB 3406 GGTGGGCAACACAT 3420

RESULT 47  
US-10-758-672A-18  
Sequence 18, Application US/10758672A  
Publication No. US20040185037A1  
GENERAL INFORMATION:  
APPLICANT: Han, et al.  
TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
FILE REFERENCE: 01017/35966B  
CURRENT APPLICATION NUMBER: US/10/758,672A  
CURRENT FILING DATE: 2004-01-15  
PRIOR APPLICATION NUMBER: US 09/724,126  
PRIOR FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: US 60/187,911  
PRIOR FILING DATE: 2000-03-08

; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 5205  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (662)..(662)  
; OTHER INFORMATION: n = a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (668)..(668)  
; OTHER INFORMATION: n = a, c, g, or t  
US-10-758-672A-18

Query Match 75.0%; Score 15; DB 8; Length 5205;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20  
Db 1518 GGTGGGCAACACAT 1532

## RESULT 48

US-10-758-636A-18  
; Sequence 18, Application US/10758636A  
; Publication No. US20050089876A1  
; GENERAL INFORMATION:  
; APPLICANT: Han, et al.

; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY  
; FILE REFERENCE: 01017/35966C  
; CURRENT APPLICATION NUMBER: US/10/758,636A  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/1724,126  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: US 60/187,911  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 5205  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (662)..(662)  
; OTHER INFORMATION: n = a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (668)..(668)  
; OTHER INFORMATION: n = a, c, g, or t  
US-10-758-636A-18

Query Match 75.0%; Score 15; DB 9; Length 5205;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20  
Db 1518 GGTGGGCAACACAT 1532

## RESULT 49

US-10-357-819-9  
; Sequence 9, Application US/10357819  
; Publication No. US2004025974A1  
; GENERAL INFORMATION:  
; APPLICANT: Alvarez, Enrique  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Gangolli, Esna A.  
; APPLICANT: Gerlach, Valerie L.

; APPLICANT: Gorman, Linda  
; APPLICANT: Guo, Xiaolia  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Kekuda, Rameesh  
; APPLICANT: Li, Li  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Zhong, Mei

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD OF  
; FILE REFERENCE: 21402-538A

; CURRENT APPLICATION NUMBER: US/10/357,819  
; CURRENT FILING DATE: 2003-02-03  
; PRIOR APPLICATION NUMBER: 09/520,781  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 09/584,411  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 09/783,436  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 10/085,198  
; PRIOR FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/353,301  
; PRIOR FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 60/355,099  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: 60/356,424  
; PRIOR FILING DATE: 2002-02-12  
; PRIOR APPLICATION NUMBER: 60/358,239  
; PRIOR FILING DATE: 2002-02-20  
; PRIOR APPLICATION NUMBER: 60/358,608  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: 60/359,367  
; PRIOR FILING DATE: 2002-02-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 9  
; LENGTH: 5466  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (66)..(5312)  
US-10-357-819-9

Query Match 75.0%; Score 15; DB 8; Length 5466;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTGGGCAACACAT 20  
Db 1628 GGTGGGCAACACAT 1642

## RESULT 50

US-11-097-143-23077/c  
; Sequence 23077, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; FILE REFERENCE: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832



```
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23077
; LENGTH: 5492
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-23077

Query Match      75.0%; Score 15; DB 10; Length 5492;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CTTGGTTGGGCAC 16
Db      4201 CTTGGTTGGGCAC 4187
```

Search completed: May 7, 2006, 23:36:31  
Job time : 1085 secs

THIS PAGE LEFT BLANK